

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2004年12月2日 (02.12.2004)

PCT

(10) 国際公開番号
WO 2004/104986 A1

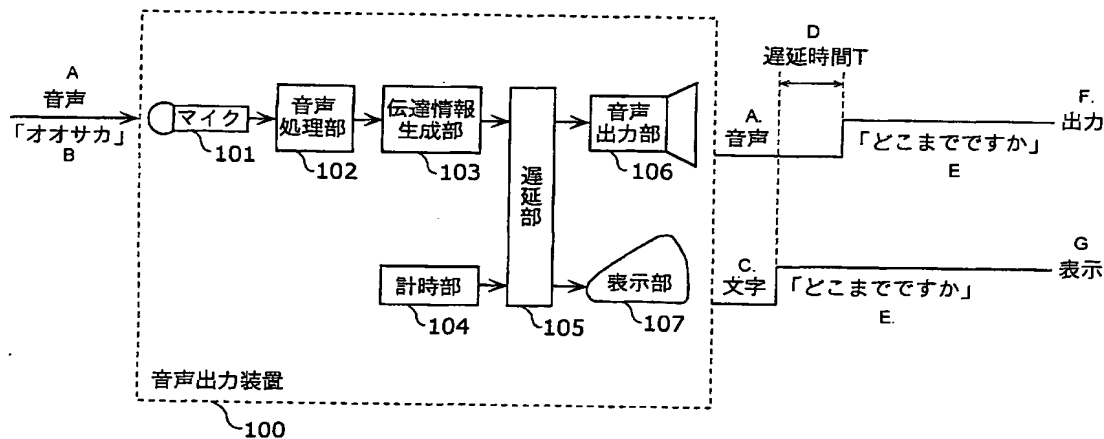
- (51) 国際特許分類⁷: G10L 13/00
(21) 国際出願番号: PCT/JP2004/006065
(22) 国際出願日: 2004年4月27日 (27.04.2004)
(25) 国際出願の言語: 日本語
(26) 国際公開の言語: 日本語
(30) 優先権データ:
特願2003-143043 2003年5月21日 (21.05.2003) JP
(71) 出願人 (米国を除く全ての指定国について): 松下電器産業株式会社 (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) [JP/JP]; 〒5718501 大阪府門真市大字門真1006番地 Osaka (JP).
(72) 発明者; および
(75) 発明者/出願人 (米国についてのみ): 西崎 誠 (NISHIZAKI, Makoto). 小沼 知浩 (KONUMA, Tomohiro). 遠藤 充 (ENDO, Mitsuru).

- (74) 代理人: 新居 広守 (NII, Hiromori); 〒5320011 大阪府大阪市淀川区西中島3丁目11番26号 新大阪末広センタービル3F 新居国際特許事務所内 Osaka (JP).
(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF,

[続葉有]

(54) Title: VOICE OUTPUT DEVICE AND VOICE OUTPUT METHOD

(54) 発明の名称: 音声出力装置及び音声出力方法



A...VOICE
B..."OSAKA"
100...VOICE OUTPUT DEVICE
101...MICROPHONE
102...VOICE PROCESSING PART
103...COMMUNICATION INFORMATION PRODUCING PART
104...TIMING PART
105...DELAY PART
106...VOICE OUTPUT PART
107...DISPLAY PART
C...TEXT
D...DELAY TIME T
E..."WHERE TO?"
F...OUTPUT
G...DISPLAY

(57) Abstract: A voice output device for ensuring that text and voice information are conveyed to the user so as to improve reliability of the interface between the device and the user, comprising a display part (107) for displaying, in text, device communication information to be conveyed to the user; and delay (105) and voice output (106) parts for estimating a delay time (T) required for the user to start reading and understand the text displayed on the display part (107) and for outputting, in voice, the device communication information when the estimated delay time (T) has elapsed since the appearance of the text for display.

[続葉有]



BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, — 補正書・説明書
TD, TG).

添付公開書類:
— 国際調査報告書

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

(57) 要約: 文字と音声による情報をユーザに対して確実に伝えてユーザとの間のインターフェースの頑健性を向上する音声出力装置は、ユーザに対して伝達すべき装置伝達情報を文字で表示する表示部(107)と、ユーザがその表示部(107)で表示される文字を視認するための動作に要する遅延時間(T)を推定し、その文字が表示されてから遅延時間(T)が経過したときに、その装置伝達情報を音声で出力する遅延部(105)及び音声出力部(106)とを備える。

明 細 書

音声出力装置及び音声出力方法

5 技術分野

本発明は、ユーザに対して音声で情報を伝える音声出力装置に関し、特に、音声を出力するとともにその音声と同一の内容を示す文字を表示する音声出力装置に関する。

10 背景技術

従来より、ユーザに対して音声で情報を伝える音声出力装置が提供され、この音声出力装置は、カーナビゲーションシステムの端末や、テレビ、パーソナルコンピュータなどのインターフェースに適用されている。

また、このような音声出力装置には、ユーザに対して情報を確実に伝えるため、音声を出力するのみならず、その情報を文字で表示するものもある（特開平 1 1 - 1 4 5 9 5 5 号公報、特開平 1 1 - 3 3 9 0 5 8 号公報、特開 2 0 0 1 - 1 4 2 4 8 4 号公報、及び特開平 5 - 2 1 6 6 1 8 号公報参照）。仮にユーザが音声を聞き逃しても、ユーザは音声出力装置をわざわざ操作することなく、表示された文字を読むことで音声出力装置から伝えられる情報を把握することができる。

図 1 は、音声及び文字で情報を伝える従来の音声出力装置の構成を示す構成図である。

この音声出力装置 9 0 0 は、対話形式でユーザから必要な情報を取得して、そのユーザの要望する鉄道の切符を販売するものであって、マイク 9 0 1 と、音声処理部 9 0 2 と、伝達情報生成部 9 0 3 と、音声出力部 9 0 4 と、表示部 9 0 5 とを備える。

マイク 901 は、ユーザからの音声を取得する。

音声処理部 902 は、マイク 901 で取得された音声から、ユーザが音声出力装置 900 に対して伝えようとするユーザ伝達情報を特定し、そのユーザ伝達情報を伝達情報生成部 903 に出力する。例えば、ユーザがマイク 901 に対して「オオサカ」と発すると、音声処理部 902 は、駅名の「大阪駅」をユーザ伝達情報として特定する。

伝達情報生成部 903 は、音声処理部 902 で特定されたユーザ伝達情報に基づいて、ユーザに対して伝えるべき装置伝達情報を生成し、その装置伝達情報を音声出力部 904 及び表示部 905 に出力する。例えば、ユーザ伝達情報が出発駅の「大阪駅」を示す場合には、伝達情報生成部 903 は、到着駅を尋ねる内容の装置伝達情報を生成して、その装置伝達情報を出力する。

音声出力部 904 は、伝達情報生成部 903 から装置伝達情報を取得すると、その装置伝達情報の内容を音声で出力する。例えば、音声出力部 904 は、到着駅を尋ねる内容の装置伝達情報を取得すると、「どこまでですか」という音声で出力する。

表示部 905 は、伝達情報生成部 903 から装置伝達情報を取得すると、その装置伝達情報の内容を文字で表示する。例えば、表示部 905 は、到着駅を尋ねる内容の装置伝達情報を取得すると、「どこまでですか」という文字を表示する。

図 2 は、音声出力装置 900 の表示部 905 が表示する画面の一例を示す画面表示図である。

表示部 905 は、条件欄 905a と指定欄 905b と質問欄 905c とを表示する。条件欄 905a には、ユーザに対して出発駅や到着駅などの問い合わせるべき内容が表示され、指定欄 905b には、ユーザから伝えられた駅名などが表示され、質問欄 905c には上述の装置伝達

情報の内容が文字で表示される。

ユーザは、このような音声出力装置 900 を対話形式で操作することにより所望の切符を購入する。

ここで、従来の音声出力装置 900 は、音声の出力と文字の表示とを
5 同時に行う（特開平 5-216618 号公報参照）。例えば、音声出力部 904 が「どこまでですか」という音声を出力すると同時に、表示部 905 が「どこまでですか」という文字を表示する。

しかしながら、上記従来の音声出力装置 900 では、音声は文字の表示と同時に出力されるため、ユーザの注意力が文字よりも音声に集中してしまっ
10 てしまって、文字表示がユーザにとって無意味となり、ユーザとの間のインターフェースの頑健性を向上することができないという問題がある。

これは、人は文字が表示されてもその文字を認識するのに時間を要してしまうからである。人は、文字が表示されてから眼球運動を開始する
15 までに、70ms から 700ms までの時間を要することが知られている（田村博著「ヒューマンインターフェース」（1998 年オーム社発行）参照）。また、その時間の平均は 200ms である。更に、視点を文字の位置まで動かし、焦点をその文字に合わせるまでには、それ以上の時間が必要とされる。

20 本発明は、かかる問題に鑑みてなされたものであり、文字と音声による情報をユーザに対して確実に伝えてユーザとの間のインターフェースの頑健性を向上する音声出力装置を提供することを目的とする。

発明の開示

25 上記目的を達成するために、本発明の音声出力装置は、ユーザに対して伝達すべき伝達情報を文字で表示する文字表示手段と、前記文字表示

手段に文字が表示されてから、ユーザが前記文字を視認するための動作に要する遅延時間が経過したときに、前記伝達情報を音声で出力する音声出力手段とを備えることを特徴とする。

- これにより、伝達情報を示す文字が表示されてから遅延時間経過後に
- 5 その伝達情報を示す音声出力されるため、ユーザは、眼球を動かして表示された文字に焦点を合わせた状態から、その文字の認識と、音声の認識とを同時に開始し、音声と文字の両方に注意を払うことができる。その結果、文字と音声による情報をユーザに対して確実に伝えてユーザとの間のインターフェースの頑健性を向上することができる。
- 10 また、前記音声出力装置は、さらに、前記文字表示手段に表示される文字の表示態様に応じて前記遅延時間を推定する遅延推定手段を備え、前記音声出力手段は、前記文字表示手段に文字が表示されてから、前記遅延推定手段により推定された遅延時間が経過したときに、前記伝達情報を音声で出力することを特徴としても良い。例えば、前記遅延推定手段は、前記文字表示手段により文字が表示されてから、ユーザの視点が
- 15 前記文字に移動を開始するまでの開始時間を含むように前記遅延時間を推定する。また、前記遅延推定手段は、さらに、ユーザの視点が移動を開始してから前記文字に到達するまでの移動時間を含むように前記遅延時間を推定する。また、前記遅延推定手段は、さらに、ユーザの視点が
- 20 前記文字に到達してから前記文字に焦点が合うまでの焦点合わせ時間を含むように前記遅延時間を推定する。

これにより、文字表示手段に表示される文字の表示態様が異なっても、その文字の表示態様に応じて遅延時間が推定されるため、ユーザはその文字の認識と音声の認識とを同時に開始して、音声と文字の両方に注意

25 を払うことができる。

また、前記音声出力装置は、さらに、ユーザの特徴を示す個人情報を

取得する個人情報取得手段を備え、前記遅延推定手段は、前記個人情報取得手段により取得された個人情報に基づいて、前記ユーザに応じた遅延時間を推定することを特徴としても良い。例えば、前記個人情報取得手段は、前記ユーザの年齢を前記個人情報として取得し、前記遅延推定手段は、前記個人情報取得手段により取得された年齢に基づいて、前記ユーザに応じた遅延時間を推定する。

これにより、遅延時間がユーザの特徴を示す年齢に基づいて推定されるため、個々のユーザの何れに対しても、そのユーザの年齢に応じた遅延時間だけ音声の出力を文字表示から遅らせることができ、文字と音声による情報をユーザに対してより確実に伝えることができる。

また、前記音声出力装置は、さらに、ユーザによる操作に応じて、前記文字表示手段に文字を表示させるとともに、前記音声出力手段に音声を出力させる操作手段と、ユーザの前記操作手段に対する操作の慣れ度合いを特定する慣れ特定手段とを備え、前記遅延推定手段は、前記慣れ特定手段により特定された慣れ度合いに基づいて、前記ユーザの慣れに応じた遅延時間を推定することを特徴としても良い。例えば、前記慣れ特定手段は、前記操作手段に対するユーザの操作回数を前記慣れ度合いとして特定する。

これにより、遅延時間がユーザの慣れ度合いに基づいて推定されるため、操作手段を操作することによりユーザの慣れ度合いが変化しても、その慣れ度合いに応じた遅延時間だけ音声の出力を文字表示から遅らせることができ、文字と音声による情報をユーザに対してより確実に伝えることができる。

また、前記遅延推定手段は、ユーザの注意を引き付ける前記音声出力装置の注視点から、前記文字表示手段により表示された文字までの文字表示距離に基づいて、前記焦点合わせ時間を特定することを特徴として

も良い。

通常、文字表示距離が短ければ焦点合わせ時間も短く、文字表示距離が長ければ焦点合わせ時間も長いため、このように焦点合わせ時間を文字表示距離に基づいて特定することにより、適切な焦点合わせ時間を特定することができる。

また、前記遅延推定手段は、シグモイド関数を利用して前記遅延時間を推定することを特徴としても良い。

シグモイド関数は生態系のモデルを表現し得るものであるため、このようにシグモイド関数を利用して遅延時間を推定することにより、生体特性に合致した適切な遅延時間を推定することができる。

また、前記遅延推定手段は、前記文字表示手段により表示される文字のサイズに基づいて、前記開始時間を特定することを特徴としても良い。

通常、文字のサイズが小さければ開始時間は長く、文字のサイズが大きければ開始時間は短くなるため、このように開始時間を文字のサイズに基づいて特定することにより、適切な開始時間を特定することができる。

なお、本発明は、上記音声出力装置によって音声が出力される音声出力方法やそのプログラムとして実現することもできる。

20 図面の簡単な説明

図1は、音声及び文字で情報を伝える従来の音声出力装置の構成を示す構成図である。

図2は、同上の表示部が表示する画面の一例を示す画面表示図である。

図3は、実施の形態における音声出力装置の構成を示す構成図である。

図4は、同上の音声出力装置の表示部が表示する画面の一例を示す画面表示図である。

図 5 は、同上の音声出力装置の動作を示すフロー図である。

図 6 は、同上の関数 $f_0(X)$ 及び運動開始時間 T_a と文字サイズ X との関係を示す図である。

図 7 は、同上の変数 S の値によって変化する関数 $f_1(X)$ を示す図
5 である。

図 8 は、同上の関数 $f_2(L)$ 及び移動時間 T_b と文字表示距離 L との関係を示す図である。

図 9 は、同上の関数 $f_2(L)$ 及び焦点合わせ時間 T_c と文字表示距離 L との関係を示す図である。

10 図 10 は、同上の変形例 1 にかかる音声出力装置の構成を示す構成図である。

図 11 は、同上の変形例 1 の関数 $f_3(M)$ 及び個人別遅延時間 T_1 と年齢 M との関係を示す図である。

図 12 は、同上の変形例 2 にかかる音声出力装置の構成を示す構成図
15 である。

図 13 は、同上の変形例 2 の関数 $f_4(K)$ 及び慣れ遅延時間 T_2 と操作回数 K との関係を示す図である。

発明を実施するための最良の形態

20 以下、本発明の実施の形態における音声出力装置について、図面を参照しながら説明する。

図 3 は、本実施の形態における音声出力装置の構成を示す構成図である。

本実施の形態における音声出力装置 100 は、ユーザに伝える情報を
25 音声で出力するとともにその情報を出す文字を表示するものであって、
マイク 101 と、音声処理部 102 と、伝達情報生成部 103 と、計時

部 1 0 4 と、遅延部 1 0 5 と、音声出力部 1 0 6 と、表示部 1 0 7 とを備えている。

このような音声出力装置 1 0 0 は、音声の出力時刻を、文字の表示時刻から、人がその文字を視認するための動作に要する時間（以下、遅延
5 時間という）だけ遅らせることにより、その音声と文字とをユーザに対して確実に認識させる点に特徴がある。

マイク 1 0 1 は、ユーザからの音声を取得する。

音声処理部 1 0 2 は、マイク 1 0 1 で取得された音声から、ユーザが音声出力装置 1 0 0 に対して伝えようとするユーザ伝達情報を特定し、
10 そのユーザ伝達情報を伝達情報生成部 1 0 3 に出力する。例えば、ユーザがマイク 1 0 1 に対して「オオサカ」と発すると、音声処理部 1 0 2 は、駅名の「大阪駅」をユーザ伝達情報として特定する。

伝達情報生成部 1 0 3 は、音声処理部 1 0 2 で特定されたユーザ伝達情報に基づいて、ユーザに対して伝えるべき装置伝達情報を生成し、その装置伝達情報を遅延部 1 0 5 に出力する。例えば、ユーザ伝達情報が
15 出発駅の「大阪駅」を示す場合には、伝達情報生成部 1 0 3 は、到着駅を尋ねる内容の装置伝達情報を生成して、その装置伝達情報を出力する。

計時部 1 0 4 は、遅延部 1 0 5 からの指示に応じて時間を計測し、その計測結果を遅延部 1 0 5 に出力する。

遅延部 1 0 5 は、伝達情報生成部 1 0 3 から装置伝達情報を取得すると、その装置伝達情報を表示部 1 0 7 に対して出力するとともに、計時部 1 0 4 に時間の計測を開始させる。そして、遅延部 1 0 5 は、表示部 1 0 7 に表示される文字の表示態様に応じて上述の遅延時間を推定し、計時部 1 0 4 で計測された計測時間が遅延時間に達したときに、装置伝
25 達情報を音声出力部 1 0 6 に出力する。

表示部 1 0 7 は、遅延部 1 0 5 から装置伝達情報を取得すると、その

装置伝達情報の内容を文字で表示する。例えば、表示部 107 は、到着駅を尋ねる内容の装置伝達情報を取得すると、「どこまでですか」という文字を表示する。

音声出力部 106 は、遅延部 105 から装置伝達情報を取得すると、
5 その装置伝達情報の内容を音声で出力する。例えば、音声出力部 106 は、到着駅を尋ねる内容の装置伝達情報を取得すると、「どこまでですか」という音声を出力する。

図 4 は、音声出力装置 100 の表示部 107 が表示する画面の一例を示す画面表示図である。

10 表示部 107 は、条件欄 107a と、指定欄 107b と、質問欄 107c と、エージェント 107d と、開始ボタン 107e と、確認ボタン 107f とを表示する。

条件欄 107a には、ユーザに対して出発駅や到着駅などの問い合わせるべき内容が表示され、指定欄 107b には、ユーザから伝えられた
15 駅名などが表示され、質問欄 107c には上述の装置伝達情報の内容が文字で表示される。また、質問欄 107c の文字は、あたかもエージェント 107d が喋っているように表示される。

開始ボタン 107e は、ユーザに選択されることにより、音声出力装置 100 の対話形式による切符販売動作を開始させる。

20 確認ボタン 107f は、ユーザに選択されることにより、ユーザから取得された出発駅や到着駅などの情報に応じた切符の発券を開始する。

図 5 は、音声出力装置 100 の動作を示すフロー図である。

まず、音声出力装置 100 は、ユーザからの音声を取得し（ステップ S100）、その取得した音声からユーザ伝達情報を特定する（ステップ S102）。
25

次に、音声出力装置 100 は、そのユーザ伝達情報に基づいて、その

情報に対応する装置伝達情報を生成し（ステップS104）、その装置伝達情報を文字で表示するとともに（ステップS106）、時間の計測を開始する（ステップS108）。

このように時間の計測を開始すると、音声出力装置100は、その文字の表示態様を考慮して遅延時間Tを推定し、計測時間がその遅延時間T以上であるか否かを判別する（ステップS110）。ここで、音声出力装置100は、遅延時間T未満であると判別すると（ステップS110のNo）、ステップS108からの動作を繰り返し実行する。即ち、時間計測を継続して実行する。一方、音声出力装置100は、遅延時間T以上であると判別すると（ステップS110のYes）、装置伝達情報を音声で出力する（ステップS112）。

ここで、遅延部105は、表示部107に表示される文字の表示態様に応じ、運動開始時間Taと移動時間Tbと焦点合わせ時間Tcとを考慮して、上述の遅延時間Tを推定する。

運動開始時間Taは、文字が表示されてから、ユーザの視点がその文字に向かって動き出すまでに要する時間である。例えば、ユーザが表示部107のエージェント107dを注視しているときに、「どこまでですか」という文字が質問欄107cに表示される場合、運動開始時間Taは、その文字が表示されてから、ユーザが注視点であるエージェント107dから視点を外すまでに要する時間である。

移動時間Tbは、ユーザの視点が文字に向かって動き出してから、その文字に到達するまでに要する時間である。例えば、ユーザが注視しているエージェント107dから、質問欄107cの文字までの距離が長いと、当然に視点を動かす距離も長くなり、その結果、移動時間Tbも長くなる。このような場合には、遅延時間Tはその移動時間Tbを考慮して決定される必要がある。

焦点合わせ時間 T_c は、ユーザの視点が文字に到達してから、その文字に焦点が合うまでに要する時間である。一般に、人が注視しているものから別のものを見るために視点を移動すると、その移動距離が長いほど焦点のずれが生じる。そこで、このような焦点合わせ時間 T_c は、視
 5 点の移動距離に応じて特定される。

ここで、運動開始時間 T_a について詳細に説明する。

この運動開始時間 T_a は、表示される文字のサイズによって変化する。文字のサイズが大きくなれば、ユーザの注意はその文字に強く引き付けられて、運動開始時間 T_a は短くなる。一方、文字のサイズが小さくな
 10 れば、ユーザの注意を文字に引き付ける力は弱く、運動開始時間 T_a は長くなる。例えば、文字の標準のサイズを 10 ポイントとすると、文字のサイズを 10 ポイントより大きくするほど、ユーザの注意を文字に引き付ける力が大きくなり、運動開始時間 T_a は短くなる。

遅延部 105 は、以下の（式 1）から運動開始時間 T_a を導出する。
 15
$$T_a = t_0 - \alpha_0 \dots \text{（式 1）}$$

t_0 は、文字のサイズを限りなく小さくしたときに要する一定の時間である。運動開始時間 T_a は、この時間 t_0 から、文字のサイズによって変化する時間 α_0 を減算することによって導出される。

遅延部 105 は、以下の（式 2）から時間 α_0 を導出する。
 20
$$\alpha_0 = t_1 * f_0(X) \dots \text{（式 2）}$$

X は文字サイズを示し、 t_1 は文字サイズ X によって短縮可能な最大の時間である。なお、記号「 $*$ 」は積を示す。

また、関数 $f_0(X)$ は以下の（式 3）によって示される。

$$f_0(X) = 1 / (1 + \exp(-((X - X_A) / (X_C - X_A) - 0.5) / 0.1)) \dots \text{（式 3）}$$

 25

X_A は運動開始時間 T_a を決めるための基準の文字サイズ（例えば、

10 ポイント) であり、 X_C は運動開始時間 T_a を決めるための最大の文字サイズ (例えば、38 ポイント) である。なお、記号「exp」は自然対数の底を示し、 $\exp(A)$ は、自然対数の底の A 乗を示す。

5 このような関数 $f_0(X)$ は、生態系のモデルとして良く利用されるシグモイド関数である。即ち、このような関数 $f_0(X)$ を用いることで、文字サイズ X に応じた眼球運動特性に適合する運動開始時間 T_a を導出することができる。

図 6 は、関数 $f_0(X)$ 及び運動開始時間 T_a と文字サイズ X との関係を示す図である。

10 関数 $f_0(X)$ によって示される値は、図 6 の (a) に示すように、文字サイズ X が基準サイズ X_A から最大ポイント X_C に変化するに伴って増加する。即ち、この値は文字サイズ X の増加とともに、基準サイズ X_A (10 ポイント) 付近で緩やかに増加し、中間サイズ (24 ポイント) 付近で急激に増加し、最大サイズ X_C (38 ポイント) 付近で再び
15 緩やかに増加する。

したがって、運動開始時間 T_a は、図 6 の (b) に示すように、文字サイズ X の増加に伴って、基準サイズ X_A 付近で緩やかに減少し、中間サイズ付近で急激に減少し、最大サイズ X_C 付近で再び緩やかに減少する。

20 ここで、関数 $f_0(X)$ の代わりに (式 4) に示す関数を用いても良い。

$$f_1(X) = 1 / (1 + \exp(-S * ((X - X_A) / (X_C - X_A) - 0.5) / 0.1)) \dots \text{(式 4)}$$

S はシグモイド関数の変極点の傾きを決定する変数である。

25 図 7 は、変数 S によって変化する関数 $f_1(X)$ を示す図である。

この図 7 に示すように、関数 $f_1(X)$ によって示される値は、変数

S を小さくすると、変極点（中間サイズ）付近で緩やかに変化するが、変数 S を大きくすると、変極点付近で急激に変化する。この変数 S を適当な値に設定することにより、より正確な運動開始時間 T_a を導出することができる。

5 次に、移動時間 T_b について詳細に説明する。

この移動時間 T_b は、注視点であるエージェント 107d から質問欄 107c の文字までの距離（以下、文字表示距離という）によって定められる。

遅延部 105 は、以下の（式 5）から移動時間 T_b を導出する。

10
$$T_b = t_0 + \alpha_1 \cdot \cdot \cdot \text{（式 5）}$$

t_0 は、文字表示距離が 0 のときに要する一定の時間である。移動時間 T_b は、この時間 t_0 に対して、文字表示距離に応じて変化する時間 α_1 を加算することによって導出される。

遅延部 105 は、以下の（式 6）から時間 α_1 を導出する。

15
$$\alpha_1 = t_2 * f_2(L) \cdot \cdot \cdot \text{（式 6）}$$

L は文字表示距離であり、 t_2 は文字表示距離 L によって延長され得る最大の時間である。

また、関数 $f_2(L)$ は以下の（式 7）によって示される。

20
$$f_2(L) = 1 / (1 + \exp(-(L - L_A) / (L_C - L_A) - 0.5) / 0.1)) \cdot \cdot \cdot \text{（式 7）}$$

L_A は基準距離を示し、 L_C は最大距離を示す。例えば、基準距離は 0 cm であり、最大距離は 10 cm である。

このような関数 $f_2(L)$ は、生態系のモデルとして良く利用されるシグモイド関数である。即ち、このような関数 $f_2(L)$ を用いることで、文字表示距離 L に応じた眼球運動特性に適合する移動時間 T_b を導出することができる。

25

また、文字表示距離 L は以下の (式 8) によって示される。

$$L = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2} \quad \dots (式 8)$$

p_x 及び p_y は、それぞれ質問欄 107c の文字の X 座標位置及び Y 座標位置を示し、 q_x 及び q_y は、それぞれエージェント 107d の X 座標位置及び Y 座標位置を示す。なお、記号「 $\sqrt{}$ 」は根を示し、 \sqrt{A} は A の根を示す。また、記号「 $^{}$ 」はべき乗を示し、 A^B は A の B 乗を示す。

図 8 は、関数 $f_2(L)$ 及び移動時間 T_b と文字表示距離 L との関係を示す図である。

関数 $f_2(L)$ によって示される値は、図 8 の (a) に示すように、文字表示距離 L が基準距離 L_A から最大距離 L_C に変化するに伴って増加する。即ち、この値は文字表示距離 L の増加とともに、基準距離 L_A (0 cm) 付近で緩やかに増加し、中間距離 (5 cm) 付近で急激に増加し、最大距離 L_C (10 cm) 付近で再び緩やかに増加する。

したがって、移動時間 T_b は、図 8 の (b) に示すように、文字表示距離 L の増加に伴って、基準距離 L_A 付近で緩やかに増加し、中間距離付近で急激に増加し、最大距離 L_C 付近で再び緩やかに増加する。

なお、上述では文字表示距離 L を、エージェント 107d の位置から質問欄 107c の文字までの距離としたが、エージェント 107d が表示されていない場合には、画面の中央を注視点として、その中央から文字までの距離としても良い。

次に、焦点合わせ時間 T_c について詳細に説明する。

この焦点合わせ時間 T_c は、移動距離 T_b と同様、文字表示距離 L によって定められる。

遅延部 105 は、以下の (式 9) から焦点合わせ時間 T_c を導出する。

$$T_c = t_0 + \alpha L^2 \quad \dots (式 9)$$

t_0 は、文字表示距離 L が 0 のときに要する一定の時間である。焦点合わせ時間 T_c は、この時間 t_0 に対して、文字表示距離 L に応じて変化する時間 α_2 を加算することによって導出される。

遅延部 105 は、以下の (式 10) から時間 α_2 を導出する。

5 $\alpha_2 = t_3 * f_2(L) \cdots \text{(式 10)}$

t_3 は、文字表示距離 L によって延長され得る最大の時間である。関数 $f_2(L)$ は上述の (式 7) によって示される。

このような関数 $f_2(L)$ を用いることで、文字表示距離 L に応じた眼球運動特性に適合する焦点合わせ時間 T_c を導出することができる。

10 図 9 は、関数 $f_2(L)$ 及び焦点合わせ時間 T_c と文字表示距離 L との関係を示す図である。

関数 $f_2(L)$ によって示される値は、図 9 の (a) に示すように、文字表示距離 L が基準距離 L_A から最大距離 L_C に変化するに伴って増加する。

15 したがって、焦点合わせ時間 T_c は、図 9 の (b) に示すように、文字表示距離 L の増加に伴って、基準距離 L_A 付近で緩やかに増加し、中間距離付近で急激に増加し、最大距離 L_C 付近で再び緩やかに増加する。

遅延部 105 は、上述のような運動開始時間 T_a と、移動時間 T_b と、焦点合わせ時間 T_c とを考慮して、以下に示す (式 11) から遅延時間 T を導出する。

20 $T = t_0 - \alpha_0 + \alpha_1 + \alpha_2 \cdots \text{(式 11)}$

このように、運動開始時間 T_a と移動時間 T_b と焦点合わせ時間 T_c とを考慮して遅延時間 T が導出されることにより、遅延時間 T を人の眼球の動きに応じた正確な時間とすることができる。

25 このように本実施の形態では、ユーザが文字を視認するための動作に要する遅延時間 T を推定し、文字を表示させてから遅延時間 T だけ経過

した後に、音声を出力するため、ユーザは、文字の認識を開始すると同時に音声の認識も開始することができる。その結果、文字と音声による情報をユーザに対して確実に伝えてユーザとの間のインターフェースの頑健性を向上することができる。

- 5 ここで、PDA (personal digital assistant) などの携帯端末が備えるディスプレイのように、表示部 107 の表示画面が小さい場合には、文字表示距離 L に関わりなく時間 $\alpha 1$ と時間 $\alpha 2$ をそれぞれ一定の時間としても良い。即ち、時間 $\alpha 1$ 及び時間 $\alpha 2$ を、文字表示距離 L の変化に応じて取り得る平均的な時間とする。このように、時間 $\alpha 1$ 及び時間
- 10 $\alpha 2$ をそれぞれ平均的な時間とすることにより、遅延部 105 は遅延時間 T を以下の (式 12) から導出する。

$$T = t_0 - \alpha_0 + \text{average}(\alpha 1) + \text{average}(\alpha 2) \cdots \text{(式 12)}$$

$\text{average}(\alpha 1)$ は時間 $\alpha 1$ の平均時間を示し、 $\text{average}(\alpha 2)$ は時間 $\alpha 2$ の平均時間を示す。

- 15 このように平均時間を用いることにより、遅延時間 T の導出のためのパラメータ数を削減して、計算処理を簡略化することができる。また、その結果、遅延時間 T の算出速度を速めることができ、さらに遅延部 105 の構成を簡略化することができる。

- また、遅延時間 T の上限を決めておくことで、遅延時間 T が大きくなりすぎることを避ける事ができる。
- 20

(変形例 1)

次に、上記本実施の形態における音声出力装置の第 1 の変形例について説明する。

- 本変形例にかかる音声出力装置は、各ユーザに応じた遅延時間を推定
- 25 するものであって、具体的には各ユーザの年齢に応じて推定する。

一般に、加齢とともに眼球を動かすタイミングや移動速度、焦点を合

わせるスピードは遅くなるため、運動開始時間 T_a と移動時間 T_b と焦点合わせ時間 T_c も加齢とともに長くなる。そこで本変形例にかかる音声出力装置は、ユーザの特徴を示す年齢を考慮して遅延時間を推定する。

図 10 は、変形例 1 にかかる音声出力装置の構成を示す構成図である。

5 変形例 1 にかかる音声出力装置 100a は、マイク 101 と、音声処理部 102 と、伝達情報生成部 103 と、計時部 104 と、遅延部 105a と、音声出力部 106 と、表示部 107 と、カードリーダー 109 と、個人情報蓄積部 108 とを備える。

10 カードリーダー 109 は、ユーザによって音声出力装置 100a に挿入されるカード 109a から、個人情報である年齢や生年月日を読み出し、その読み出した個人情報を個人情報蓄積部 108 に格納する。

遅延部 105a は、まず、上述のように運動開始時間 T_a と移動時間 T_b と焦点合わせ時間 T_c とを考慮した遅延時間 T を導出する。そして遅延部 105a は、個人情報蓄積部 108 に格納されている個人情報を参照し、遅延時間 T からその個人情報を考慮した個人別遅延時間 T_1 を導出する。さらに、遅延部 105a は、装置伝達情報を音声出力部 106 から音声で出力させてから、その個人別遅延時間 T_1 の経過後に、その装置伝達情報を表示部 107 に文字で表示させる。

20 遅延部 105 は、以下の（式 13）から個人別遅延時間 T_1 を導出する。

$$T_1 = T + \alpha_3 \dots \text{（式 13）}$$

個人別遅延時間 T_1 は、遅延時間 T に対して、年齢に応じて変化する時間 α_3 を加算することによって導出される。

遅延部 105 は、以下の（式 14）から時間 α_3 を導出する。

25 $\alpha_3 = t_4 * f_3(M) \dots \text{（式 14）}$

M は年齢であり、 t_4 は年齢 M によって延長され得る最大の時間であ

る。

また、関数 $f_3(M)$ は以下の (式 15) によって示される。

$$f_3(M) = 1 / (1 + \exp(-(M - 20) / (60 - 20) - 0.5) / 0.1)) \cdots \text{(式 15)}$$

- 5 このような関数 $f_3(M)$ で示される値は加齢とともに増加するため、その増加に伴って個人別遅延時間 T_1 も増加する。

図 11 は、関数 $f_3(M)$ 及び個人別遅延時間 T_1 と年齢 M との関係を示す図である。

- 関数 $f_3(M)$ によって示される値は、図 11 の (a) に示すように、
10 年齢 M が 20 歳から 60 歳に変化するに伴って増加する。即ち、この値は加齢とともに、運動能力が活発な 20 歳 (基準年齢) 付近で緩やかに増加し、40 歳 (中間年齢) 付近で急激に増加し、運動能力が衰えた 60 歳 (最大年齢) 付近で再び緩やかに増加する。

- したがって、個人別遅延時間 T_1 は、図 11 の (b) に示すように、
15 加齢に伴って、基準年齢付近で緩やかに増加し、中間年齢付近で急激に増加し、最大年齢付近で再び緩やかに増加する。

- このように本変形例では、ユーザの年齢などを考慮して個人別遅延時間 T_1 を導出し、文字を表示してからその個人別遅延時間 T_1 だけ経過した後に音声を出力するため、各ユーザごとにインターフェースの頑健
20 性の向上を図ることができる。

- なお、本変形例では、個人情報としてユーザの年齢を用いたが、ユーザの反応速度や、視点移動速度、焦点合わせ速度、俊敏性、使用履歴などを用いても良い。このような場合、上述のような反応速度などの個人情報
25 がカード 109a に予め登録されており、カードリーダー 109 は、その個人情報をカード 109a から読み出して個人情報蓄積部 108 に格納する。遅延部 105a は、個人情報蓄積部 108 に格納されている

反応速度などの個人情報参照して、遅延時間 T からその反応速度などを考慮した個人別遅延時間 T_1 を導出する。

(変形例 2)

次に、上記本実施の形態における音声出力装置の第 2 の変形例について説明する。

本変形例にかかる音声出力装置は、ユーザの慣れに応じた遅延時間を推定するものであって、具体的にはユーザの操作回数に応じて推定する。

一般に、ユーザが音声出力装置を操作する回数が増えると、ユーザはその操作に慣れるため、運動開始時間 T_a と移動時間 T_b と焦点合わせ時間 T_c は短くなる。

例えば、切符の購入などの一連の動作において、ユーザと音声出力装置との対話が進むにつれて、ユーザは、文字の表示の位置やタイミングを学習する。その結果、ユーザは、対話つまり操作回数が増加することにより、文字が表示されてからその文字に焦点を合わせるまでの動作を効率良く行うことができる。そこで本変形例にかかる音声出力装置は、ユーザの操作回数を考慮して遅延時間を推定する。

図 12 は、変形例 2 にかかる音声出力装置の構成を示す構成図である。

変形例 2 にかかる音声出力装置 100b は、マイク 101 と、音声処理部 102 と、伝達情報生成部 103 と、計時部 104 と、遅延部 105b と、音声出力部 106 と、表示部 107 と、カウンタ 110 とを備える。

カウンタ 110 は、音声処理部 102 から出力されるユーザ伝達情報を取得すると、その取得回数、つまりユーザの音声出力装置 100b に対する操作回数を数えて、その操作回数を遅延部 105b に通知する。

遅延部 105b は、まず、上述のように運動開始時間 T_a と移動時間 T_b と焦点合わせ時間 T_c とを考慮した遅延時間 T を導出する。そして

遅延部 105b は、カウンタ 110 から通知される操作回数を参照し、遅延時間 T からその操作回数を考慮した慣れ遅延時間 T2 を導出する。さらに、遅延部 105b は、装置伝達情報を音声出力部 106 から音声で出力させてから、その慣れ遅延時間 T2 の経過後に、その装置伝達情報
5 報を表示部 107 に文字で表示させる。

遅延部 105b は、以下の (式 16) から慣れ遅延時間 T2 を導出する。

$$T2 = T - \alpha 4 \cdots \text{(式 16)}$$

慣れ遅延時間 T2 は、遅延時間 T から、操作回数に応じて変化する時間 $\alpha 4$ を減算することによって導出される。
10

遅延部 105b は、以下の (式 17) から時間 $\alpha 4$ を導出する。

$$\alpha 4 = t 5 * f 4 (K) \cdots \text{(式 17)}$$

K は操作回数であり、t 5 は操作回数 K によって短縮可能な最大の時間である。

また、関数 $f 4 (K)$ は以下の (式 18) によって示される。
15

$$f 4 (K) = 1 / (1 + \exp(-(K/KC - 0.5) / 0.1)) \cdots \text{(式 18)}$$

ここで KC は、慣れ遅延時間 T2 が最短となるような操作回数の最大値である。

このような関数 $f 4 (K)$ で示される値は操作回数 K の増加に伴って増加するため、慣れ遅延時間 T2 は操作回数 K の増加に伴って減少する。
20

図 13 は、関数 $f 4 (K)$ 及び慣れ遅延時間 T2 と操作回数 K との関係を示す図である。

関数 $f 4 (K)$ によって示される値は、図 13 の (a) に示すように、操作回数 K が 0 回 (基準回数) から KC 回 (最大回数) へ変化するに伴って増加する。即ち、この値は操作回数 K の増加とともに、全く不慣れな状態の 0 回付近で緩やかに増加し、適度に慣れ始めた状態の KC / 2
25

回（中間回数）付近で急激に増加し、十分に慣れた状態のＫＣ回（最大回数）付近で再び緩やかに増加する。

したがって、慣れ遅延時間Ｔ２は、図１３の（ｂ）に示すように、操作回数Ｋの増加に伴って、基準回数付近で緩やかに減少し、中間回数付近で急激に減少し、最大回数付近で再び緩やかに減少する。

このように本変形例では、ユーザの慣れを考慮して慣れ遅延時間Ｔ２を導出し、文字を表示してからその慣れ遅延時間Ｔ２だけ経過した後に、音声を出力するため、ユーザの慣れに適したインターフェースの頑健性を保つことができる。

10 （変形例３）

次に、上記本実施の形態における音声出力装置の第３の変形例について説明する。

本変形例にかかる音声出力装置は、変形例２と同様、ユーザの慣れに応じた遅延時間を推定するものであって、具体的にはユーザの操作時間に応じて推定する。

一般に、ユーザが音声出力装置を操作する時間が長くなると、ユーザはその操作に慣れるため、運動開始時間Ｔ_aと移動時間Ｔ_bと焦点合わせ時間Ｔ_cは短くなる。そこで本変形例にかかる音声出力装置は、ユーザの操作時間を考慮して遅延時間を推定する。

20 変形例３にかかる音声出力装置は、図１２に示す変形例２の音声出力装置１００ｂと同様の構成を有するが、遅延部１０５ｂとカウンタ１１０の動作が異なる。

本変形例にかかるカウンタ１１０は、タイムカウンタとしての機能を有し、音声出力装置１００ｂとユーザとの間で対話が始まった後に音声処理部１０２から最初のユーザ伝達情報を取得すると、その取得したときからの経過時間、つまり操作時間を計測する。そしてカウンタ１１

0はその操作時間を遅延部105bに通知する。

遅延部105bは、まず、上述のように運動開始時間 T_a と移動時間 T_b と焦点合わせ時間 T_c とを考慮した遅延時間 T を導出する。そして遅延部105bは、カウンタ110から通知される操作時間を参照し、

5 遅延時間 T からその操作時間を考慮した慣れ遅延時間 T_3 を導出する。さらに、遅延部105bは、装置伝達情報を音声出力部106から音声で出力させてから、その慣れ遅延時間 T_3 の経過後に、その装置伝達情報を表示部107に文字で表示させる。

遅延部105bは、以下の(式19)から慣れ遅延時間 T_3 を導出する。

10

$$T_3 = T - \alpha_5 \cdots \text{(式19)}$$

慣れ遅延時間 T_3 は、遅延時間 T から、操作時間に応じて変化する時間 α_5 を減算することによって導出される。

遅延部105bは、以下の(式20)から時間 α_5 を導出する。

15

$$\alpha_5 = t_6 * f_5(P) \cdots \text{(式20)}$$

P は操作時間であり、 t_6 は操作時間 P によって短縮可能な最大の時間である。

また、関数 $f_5(P)$ は以下の(式21)によって示される。

$$f_5(P) = 1 / (1 + \exp(-(P/PC - 0.5)/0.1)) \cdots \text{(式21)}$$

20 ここで PC は、慣れ遅延時間 T_3 が最短となるような操作時間 P の最大値である。

このような関数 $f_5(P)$ で示される値は操作時間 P の増加に伴って増加するため、慣れ遅延時間 T_3 は操作時間 P の増加に伴って減少する。

このように本変形例では、変形例2と同様、ユーザの慣れを考慮して

25 慣れ遅延時間 T_3 を導出し、文字を表示してからその慣れ遅延時間 T_3 だけ経過した後に音声を出力するため、ユーザの慣れに適したインター

フェースの頑健性を保つことができる。

- なお、本変形例では、音声処理部 102 がユーザ伝達情報を出力したタイミング、即ちユーザが音声を発したタイミングで操作時間 P の計測を開始したが、音声出力装置 100b に電源が投入されたタイミング、
- 5 又は、開始ボタン 107f が選択されたタイミングで操作時間 P の計測を開始しても良い。

(変形例 4)

次に、本実施の形態における運動開始時間 T_a の導出方法に関する第 4 の変形例について説明する。

- 10 一般に、文字のサイズだけではなく、文字の表示位置によっても運動開始時間 T_a は変化する。つまり、表示される文字の位置が、ユーザの注視点に近ければ近いほど、ユーザはその文字に早く気づくため、運動開始時間 T_a は短くなる。

- 本変形例にかかる遅延部 105 は、文字表示距離 L に基づいて運動開始時間 T_a を以下の (式 22) から導出する。
- 15

$$T_a = t_0 + \alpha_6 \cdots \text{(式 22)}$$

t_0 は、文字表示距離 L が 0 のときに要する一定の時間である。即ち、運動開始時間 T_a は、この時間 t_0 に対して、文字表示距離 L によって変化する時間 α_6 を加算することによって導出される。

- 20 遅延部 105 は、以下の (式 23) から α_6 を導出する。

$$\alpha_6 = t_7 * f_2(L) \cdots \text{(式 23)}$$

t_7 は文字表示距離 L によって延長され得る最大の時間である。また、関数 $f_2(L)$ は (式 7) によって示される。

- このように本変形例では、文字表示距離 L を考慮した運動開始時間 T_a に基づく遅延時間 T を導出し、文字を表示してからその遅延時間 T だけ経過した後に音声を出力するため、各文字表示距離 L に適したインタ
- 25

一フェースの頑健性を保つことができる。

(変形例 5)

次に、本実施の形態における運動開始時間 T_a の導出方法に関する第 5 の変形例について説明する。

- 5 一般に、ユーザの注視点と、表示される文字の色のコントラストとが大きく異なれば、それだけユーザはその文字に早く気づくため、運動開始時間 T_a は短くなる。

本変形例にかかる遅延部 105 は、注視点と文字のコントラストに基づいて運動開始時間 T_a を以下の (式 24) から導出する。

10
$$T_a = t_0 - \alpha_7 \cdots (\text{式 24})$$

t_0 は、コントラストを限りなく小さくしたときに要する一定の時間である。即ち、運動開始時間 T_a は、この時間 t_0 から、コントラストによって変化する時間 α_7 を減算することによって導出される。

遅延部 105 は、以下の (式 25) から α_7 を導出する。

15
$$\alpha_7 = t_8 * f_6(Q) \cdots (\text{式 25})$$

Q はコントラストを示し、 t_8 はコントラスト Q によって短縮可能な最大の時間である。

また、関数 $f_6(Q)$ は以下の (式 26) によって示される。

20
$$f_6(Q) = 1 / (1 + \exp(-(Q - Q_A) / (Q_C - Q_A) - 0.5) / 0.1)) \cdots (\text{式 26})$$

Q_A は運動開始時間 T_a を決めるための基準のコントラストであり、 Q_C は運動開始時間 T_a を決めるための最大のコントラストである。

このように本変形例では、コントラストを考慮した運動開始時間 T_a に基づく遅延時間 T を導出し、文字を表示してからその遅延時間 T だけ経過した後には音声を出力するため、各コントラストに適したインターフェースの頑健性を保つことができる。

(変形例 6)

次に、本実施の形態における運動開始時間 T_a の導出方法に関する第 6 の変形例について説明する。

5 一般に、文字を赤で表示したり、点滅させたりすることにより、ユーザはその文字に早く気づくため、運動開始時間 T_a は短くなる。

本変形例にかかる遅延部 105 は、文字の表示態様の強調度合いに基づいて運動開始時間 T_a を以下の (式 27) から導出する。

$$T_a = t_0 - \alpha_8 \cdots \text{(式 27)}$$

10 t_0 は、表示形態の強調度合いを限りなく小さくしたときに要する一定の時間である。即ち、運動開始時間 T_a は、この時間 t_0 から、強調度合いによって変化する時間 α_8 を減算することによって導出される。

遅延部 105 は、以下の (式 28) から α_8 を導出する。

$$\alpha_8 = t_9 * f_7(R) \cdots \text{(式 28)}$$

15 R は強調度合いを示し、 t_9 は強調度合い R によって短縮可能な最大の時間である。

また、関数 $f_7(R)$ は以下の (式 29) によって示される。

$$f_7(R) = 1 / (1 + \exp(-(R - R_A) / (R_C - R_A) - 0.5) / 0.1)) \cdots \text{(式 29)}$$

20 R_A は運動開始時間 T_a を決めるための基準の強調度合いであり、 R_C は運動開始時間 T_a を決めるための最大の強調度合いである。

このように本変形例では、文字の強調度合いを考慮した運動開始時間 T_a に基づく遅延時間 T を導出し、文字を表示してからその遅延時間 T だけ経過した後に音声を出力するため、文字の各強調度合いに適したインターフェースの頑健性を保つことができる。

25 以上、本発明について実施の形態及び変形例を用いて説明したが、本発明はこれらに限定されるものではない。

例えば、実施の形態及び変形例では、遅延時間 T を導出するために、運動開始時間 T_a と移動時間 T_b と焦点合わせ時間 T_c とを全て考慮したが、これらのうちの少なくとも1つを考慮して遅延時間 T を導出していても良い。

- 5 また、変形例1では、ユーザの個人情報に基づいて遅延時間を導出し、変形例2では、ユーザの慣れに基づいて遅延時間を導出したが、ユーザの個人情報及び慣れの双方に基づいて遅延時間を導出していても良い。

- また、実施の形態及び変形例では、音声出力装置を音声の出力及び文字の表示により切符の販売を行う装置として説明したが、音声の出力及び文字の表示を行うものであれば他の動作を行う装置であっても良い。
- 10 例えば、テレビや、カーナビゲーションシステムの端末、携帯電話、携帯端末、パーソナルコンピュータ、電話機、ファックス、電子レンジ、冷蔵庫、掃除機、電子辞書、電子翻訳機などとして音声出力装置を構成しても良い。

15

産業上の利用の可能性

- 本発明に係る音声出力装置は、文字と音声による情報をユーザに対して確実に伝えてユーザとの間のインターフェースの頑健性を向上することができ、例えばユーザの音声に対して音声及び文字で応答することにより切符などを販売する音声応答装置などに有用である。
- 20

請 求 の 範 囲

1. ユーザに対して伝達すべき伝達情報を文字で表示する文字表示手段と、
- 5 前記文字表示手段に文字が表示されてから、ユーザが前記文字を視認するための動作に要する遅延時間が経過したときに、前記伝達情報を音声で出力する音声出力手段と
- を備えることを特徴とする音声出力装置。
- 10 2. 前記音声出力装置は、さらに、
- 前記文字表示手段に表示される文字の表示態様に応じて前記遅延時間を推定する遅延推定手段を備え、
- 前記音声出力手段は、
- 前記文字表示手段に文字が表示されてから、前記遅延推定手段により
- 15 推定された遅延時間が経過したときに、前記伝達情報を音声で出力することを特徴とする請求の範囲第1項記載の音声出力装置。
3. 前記遅延推定手段は、
- 前記文字表示手段により文字が表示されてから、ユーザの視点が前記
- 20 文字に移動を開始するまでの開始時間を含むように前記遅延時間を推定する
- ことを特徴とする請求の範囲第2項記載の音声出力装置。
4. 前記遅延推定手段は、さらに、
- 25 ユーザの視点が移動を開始してから前記文字に到達するまでの移動時間を含むように前記遅延時間を推定する

ことを特徴とする請求の範囲第3記載の音声出力装置。

5. 前記遅延推定手段は、さらに、

ユーザの視点が前記文字に到達してから前記文字に焦点が合うまでの

5 焦点合わせ時間を含むように前記遅延時間を推定する

ことを特徴とする請求の範囲第4項記載の音声出力装置。

6. 前記音声出力装置は、さらに、

ユーザの特徴を示す個人情報を取得する個人情報取得手段を備え、

10 前記遅延推定手段は、

前記個人情報取得手段により取得された個人情報に基づいて、前記ユーザに応じた遅延時間を推定する

ことを特徴とする請求の範囲第5項記載の音声出力装置。

15 7. 前記個人情報取得手段は、前記ユーザの年齢を前記個人情報として取得し、

前記遅延推定手段は、

前記個人情報取得手段により取得された年齢に基づいて、前記ユーザに応じた遅延時間を推定する

20 ことを特徴とする請求の範囲第6項記載の音声出力装置。

8. 前記個人情報取得手段は、前記ユーザの眼球を動かす眼球速度を前記個人情報として取得し、

前記遅延推定手段は、

25 前記個人情報取得手段により取得された眼球速度に基づいて、前記ユーザに応じた遅延時間を推定する

ことを特徴とする請求の範囲第 6 項記載の音声出力装置。

9. 前記音声出力装置は、さらに、

ユーザによる操作に応じて、前記文字表示手段に文字を表示させると
5とともに、前記音声出力手段に音声を出力させる操作手段と、

ユーザの前記操作手段に対する操作の慣れ度合いを特定する慣れ特定
手段とを備え、

前記遅延推定手段は、

前記慣れ特定手段により特定された慣れ度合いに基づいて、前記ユー
10ザの慣れに応じた遅延時間を推定する

ことを特徴とする請求の範囲第 5 項記載の音声出力装置。

10. 前記慣れ特定手段は、前記操作手段に対するユーザの操作回数を
前記慣れ度合いとして特定する

15 ことを特徴とする請求の範囲第 9 項記載の音声出力装置。

11. 前記慣れ特定手段は、前記操作手段に対するユーザの操作時間を
前記慣れ度合いとして特定する

ことを特徴とする請求の範囲第 9 項記載の音声出力装置。

20

12. 前記遅延推定手段は、

ユーザの注意を引き付ける前記音声出力装置の注視点から、前記文字
表示手段により表示された文字までの文字表示距離に基づいて、前記焦
点合わせ時間を特定する

25 ことを特徴とする請求の範囲第 5 項記載の音声出力装置。

13. 前記遅延推定手段は、

前記注視点と前記文字のそれぞれの位置から前記文字表示距離を導出し、前記文字表示距離に基づいて前記焦点合わせ時間を特定することを特徴とする請求の範囲第12項記載の音声出力装置。

5

14. 前記文字表示手段は、

エージェントを前記注視点として表示し、

前記遅延推定手段は、前記エージェントを起点として文字表示距離を導出する

10 ことを特徴とする請求の範囲第12項記載の音声出力装置。

15. 前記遅延推定手段は、

シグモイド関数を利用して前記遅延時間を推定する

ことを特徴とする請求の範囲第5項記載の音声出力装置。

15

16. 前記遅延推定手段は、

ユーザの注意を引き付ける前記音声出力装置の注視点から、前記文字表示手段により表示された文字までの文字表示距離に基づいて、前記移動時間を特定する

20 ことを特徴とする請求の範囲第4項記載の音声出力装置。

17. 前記遅延推定手段は、

前記文字表示手段により表示される文字のサイズに基づいて、前記開始時間を特定する

25 ことを特徴とする請求の範囲第3項記載の音声出力装置。

18. 前記遅延推定手段は、

ユーザの注意を引き付ける前記音声出力装置の注視点から、前記文字表示手段により表示された文字までの文字表示距離に基づいて、前記開始時間を特定する

5 ことを特徴とする請求の範囲第3項記載の音声出力装置。

19. 前記遅延推定手段は、

ユーザの注意を引き付ける前記音声出力装置の注視点と、前記文字表示手段により表示された文字とのコントラストに基づいて、前記開始時

10 間を特定する

ことを特徴とする請求の範囲第3項記載の音声出力装置。

20. 前記文字表示手段は、文字を点滅させて表示し、

前記遅延推定手段は、

15 前記文字表示手段により表示された文字の点滅度合いに基づいて、前記開始時間を特定する

ことを特徴とする請求の範囲第3項記載の音声出力装置。

21. 前記遅延推定手段は、

20 ユーザの視点が移動を開始してから前記文字に到達するまでの移動時間を含むように、前記遅延時間を推定する

ことを特徴とする請求の範囲第2項記載の音声出力装置。

22. 前記遅延推定手段は、

25 ユーザの視点が前記文字に到達してから前記文字に焦点が合うまでの焦点合わせ時間を含むように前記遅延時間を推定する

ことを特徴とする請求の範囲第2項記載の音声出力装置。

23. 情報処理装置が音声を出力する方法であって、

人に対して伝達すべき伝達情報を文字で表示する文字表示ステップと、

5 前記文字表示ステップで文字が表示されてから、人が前記文字を視認するための動作に要する遅延時間が経過したときに、前記伝達情報を音声で出力する音声出力ステップと

を含むことを特徴とする音声出力方法。

10 24. 前記音声出力方法は、さらに、

前記文字表示ステップで表示される文字の表示態様に応じて前記遅延時間を推定する遅延推定ステップを含み、

前記音声出力ステップでは、

15 前記文字表示ステップで文字が表示されてから、前記遅延推定ステップで推定された遅延時間が経過したときに、前記伝達情報を音声で出力する

ことを特徴とする請求の範囲第23記載の音声出力方法。

25. 前記遅延推定ステップでは、

20 前記文字表示ステップで文字が表示されてから、前記文字にユーザの視点が移動を開始するまでの開始時間を含むように前記遅延時間を推定する

ことを特徴とする請求の範囲第24項記載の音声出力方法。

25 26. 前記遅延推定ステップでは、さらに、

ユーザの視点が移動を開始してから前記文字に到達するまでの移動時

間を含むように前記遅延時間を推定する

ことを特徴とする請求の範囲第 25 項記載の音声出力方法。

27. 前記遅延推定ステップでは、さらに、

- 5 ユーザの視点が前記文字に到達してから前記文字に焦点が合うまでの
焦点合わせ時間を含むように前記遅延時間を推定する

ことを特徴とする請求の範囲第 26 項記載の音声出力方法。

28. 人に対して伝達すべき伝達情報を文字で表示する文字表示ステ
10 ップと、

前記文字表示ステップで文字が表示されてから、人が前記文字を視認
するための動作に要する遅延時間が経過したときに、前記伝達情報を音
声で出力する音声出力ステップと

をコンピュータに実行させることを特徴とするプログラム。

15

補正書の請求の範囲

[2004年9月14日(14.09.04)国際事務局受理：出願当初の請求の範囲2、24及び28は補正された；出願当初の請求の範囲1及び23は取り下げられた；他の請求の範囲は変更なし。(7頁)]

1. (削除)

5 2. (補正後) ユーザに対して伝達すべき伝達情報を文字で表示する文字表示手段と、

前記文字表示手段に文字が表示されてから、ユーザが前記文字を視認するための動作に要する遅延時間を、前記文字表示手段に表示される文字の表示態様に応じて推定する遅延推定手段と、

10 前記文字表示手段に文字が表示されてから、前記遅延推定手段により推定された遅延時間が経過したときに、前記伝達情報を音声で出力する音声出力手段と

を備えることを特徴とする音声出力装置。

15 3. 前記遅延推定手段は、

前記文字表示手段により文字が表示されてから、ユーザの視点が前記文字に移動を開始するまでの開始時間を含むように前記遅延時間を推定する

ことを特徴とする請求の範囲第2項記載の音声出力装置。

20

4. 前記遅延推定手段は、さらに、

ユーザの視点が移動を開始してから前記文字に到達するまでの移動時間を含むように前記遅延時間を推定する

ことを特徴とする請求の範囲第3記載の音声出力装置。

25

5. 前記遅延推定手段は、さらに、

ユーザの視点が前記文字に到達してから前記文字に焦点が合うまでの
焦点合わせ時間を含むように前記遅延時間を推定する

ことを特徴とする請求の範囲第4項記載の音声出力装置。

5 6. 前記音声出力装置は、さらに、

ユーザの特徴を示す個人情報を取得する個人情報取得手段を備え、
前記遅延推定手段は、

前記個人情報取得手段により取得された個人情報に基づいて、前記ユ
ーザに応じた遅延時間を推定する

10 ことを特徴とする請求の範囲第5項記載の音声出力装置。

7. 前記個人情報取得手段は、前記ユーザの年齢を前記個人情報とし
て取得し、

前記遅延推定手段は、

15 前記個人情報取得手段により取得された年齢に基づいて、前記ユーザ
に応じた遅延時間を推定する

ことを特徴とする請求の範囲第6項記載の音声出力装置。

8. 前記個人情報取得手段は、前記ユーザの眼球を動かす眼球速度を
20 前記個人情報として取得し、

前記遅延推定手段は、

前記個人情報取得手段により取得された眼球速度に基づいて、前記ユ
ーザに応じた遅延時間を推定する

ことを特徴とする請求の範囲第6項記載の音声出力装置。

25

9. 前記音声出力装置は、さらに、

ユーザによる操作に応じて、前記文字表示手段に文字を表示させるとともに、前記音声出力手段に音声を出力させる操作手段と、

ユーザの前記操作手段に対する操作の慣れ度合いを特定する慣れ特定手段とを備え、

5 前記遅延推定手段は、

前記慣れ特定手段により特定された慣れ度合いに基づいて、前記ユーザの慣れに応じた遅延時間を推定する

ことを特徴とする請求の範囲第5項記載の音声出力装置。

10 10. 前記慣れ特定手段は、前記操作手段に対するユーザの操作回数を前記慣れ度合いとして特定する

ことを特徴とする請求の範囲第9項記載の音声出力装置。

15 11. 前記慣れ特定手段は、前記操作手段に対するユーザの操作時間を前記慣れ度合いとして特定する

ことを特徴とする請求の範囲第9項記載の音声出力装置。

12. 前記遅延推定手段は、

20 ユーザの注意を引き付ける前記音声出力装置の注視点から、前記文字表示手段により表示された文字までの文字表示距離に基づいて、前記焦点合わせ時間を特定する

ことを特徴とする請求の範囲第5項記載の音声出力装置。

13. 前記遅延推定手段は、

25 前記注視点と前記文字のそれぞれの位置から前記文字表示距離を導出し、前記文字表示距離に基づいて前記焦点合わせ時間を特定する

ことを特徴とする請求の範囲第 1 2 項記載の音声出力装置。

1 4 . 前記文字表示手段は、

エージェントを前記注視点として表示し、

5 前記遅延推定手段は、前記エージェントを起点として文字表示距離を導出する

ことを特徴とする請求の範囲第 1 2 項記載の音声出力装置。

1 5 . 前記遅延推定手段は、

10 シグモイド関数を利用して前記遅延時間を推定する

ことを特徴とする請求の範囲第 5 項記載の音声出力装置。

1 6 . 前記遅延推定手段は、

15 ユーザの注意を引き付ける前記音声出力装置の注視点から、前記文字表示手段により表示された文字までの文字表示距離に基づいて、前記移動時間を特定する

ことを特徴とする請求の範囲第 4 項記載の音声出力装置。

1 7 . 前記遅延推定手段は、

20 前記文字表示手段により表示される文字のサイズに基づいて、前記開始時間を特定する

ことを特徴とする請求の範囲第 3 項記載の音声出力装置。

1 8 . 前記遅延推定手段は、

25 ユーザの注意を引き付ける前記音声出力装置の注視点から、前記文字表示手段により表示された文字までの文字表示距離に基づいて、前記開

始時間を特定する

ことを特徴とする請求の範囲第3項記載の音声出力装置。

19. 前記遅延推定手段は、

5 ユーザの注意を引き付ける前記音声出力装置の注視点と、前記文字表示手段により表示された文字とのコントラストに基づいて、前記開始時間を特定する

ことを特徴とする請求の範囲第3項記載の音声出力装置。

10 20. 前記文字表示手段は、文字を点滅させて表示し、

前記遅延推定手段は、

前記文字表示手段により表示された文字の点滅度合いに基づいて、前記開始時間を特定する

ことを特徴とする請求の範囲第3項記載の音声出力装置。

15

21. 前記遅延推定手段は、

ユーザの視点が移動を開始してから前記文字に到達するまでの移動時間を含むように、前記遅延時間を推定する

ことを特徴とする請求の範囲第2項記載の音声出力装置。

20

22. 前記遅延推定手段は、

ユーザの視点が前記文字に到達してから前記文字に焦点が合うまでの焦点合わせ時間を含むように前記遅延時間を推定する

ことを特徴とする請求の範囲第2項記載の音声出力装置。

25

23. (削除)

24. (補正後) 情報処理装置が音声を出力する方法であって、
人に対して伝達すべき伝達情報を文字で表示する文字表示ステップと、
前記文字表示ステップで文字が表示されてから、人が前記文字を視認
5 するための動作に要する遅延時間を、前記文字表示ステップで表示され
る文字の表示態様に応じて推定する遅延推定ステップと、
前記文字表示ステップで文字が表示されてから、前記遅延推定ステッ
プで推定された遅延時間が経過したときに、前記伝達情報を音声で出力
する音声出力ステップと
10 を含むことを特徴とする音声出力方法。

25. 前記遅延推定ステップでは、
前記文字表示ステップで文字が表示されてから、前記文字にユーザの
視点が移動を開始するまでの開始時間を含むように前記遅延時間を推定
15 する
ことを特徴とする請求の範囲第24項記載の音声出力方法。

26. 前記遅延推定ステップでは、さらに、
ユーザの視点が移動を開始してから前記文字に到達するまでの移動時
20 間を含むように前記遅延時間を推定する
ことを特徴とする請求の範囲第25項記載の音声出力方法。

27. 前記遅延推定ステップでは、さらに、
ユーザの視点が前記文字に到達してから前記文字に焦点が合うまでの
25 焦点合わせ時間を含むように前記遅延時間を推定する
ことを特徴とする請求の範囲第26項記載の音声出力方法。

28. (補正後) 人に対して伝達すべき伝達情報を文字で表示する文字表示ステップと、

5 前記文字表示ステップで文字が表示されてから、人が前記文字を視認するための動作に要する遅延時間を、前記文字表示ステップで表示される文字の表示態様に応じて推定する遅延推定ステップと、

前記文字表示ステップで文字が表示されてから、前記遅延推定ステップで推定された遅延時間が経過したときに、前記伝達情報を音声で出力する音声出力ステップと

10 をコンピュータに実行させることを特徴とするプログラム。

条約第 19 条（1）に基づく説明書

請求の範囲第 2 項は、補正前の請求の範囲第 1 項から形式的に独立させた。

請求の範囲第 24 項は、補正前の請求の範囲第 23 項から形式的に独立させた。

請求の範囲第 28 項は、コンピュータに実行させるステップに「前記文字表示ステップで文字が表示されてから、人が前記文字を視認するための動作に要する遅延時間を、前記文字表示ステップで表示される文字の表示態様に応じて推定する遅延推定ステップ」を追加することで、補正前の請求の範囲を減縮した。補正後の請求の範囲第 28 項は、請求の範囲第 2 項及び第 24 項の特徴と共通する。したがって、請求の範囲第 28 項に対する補正は、新規事項の追加ではない。

図1

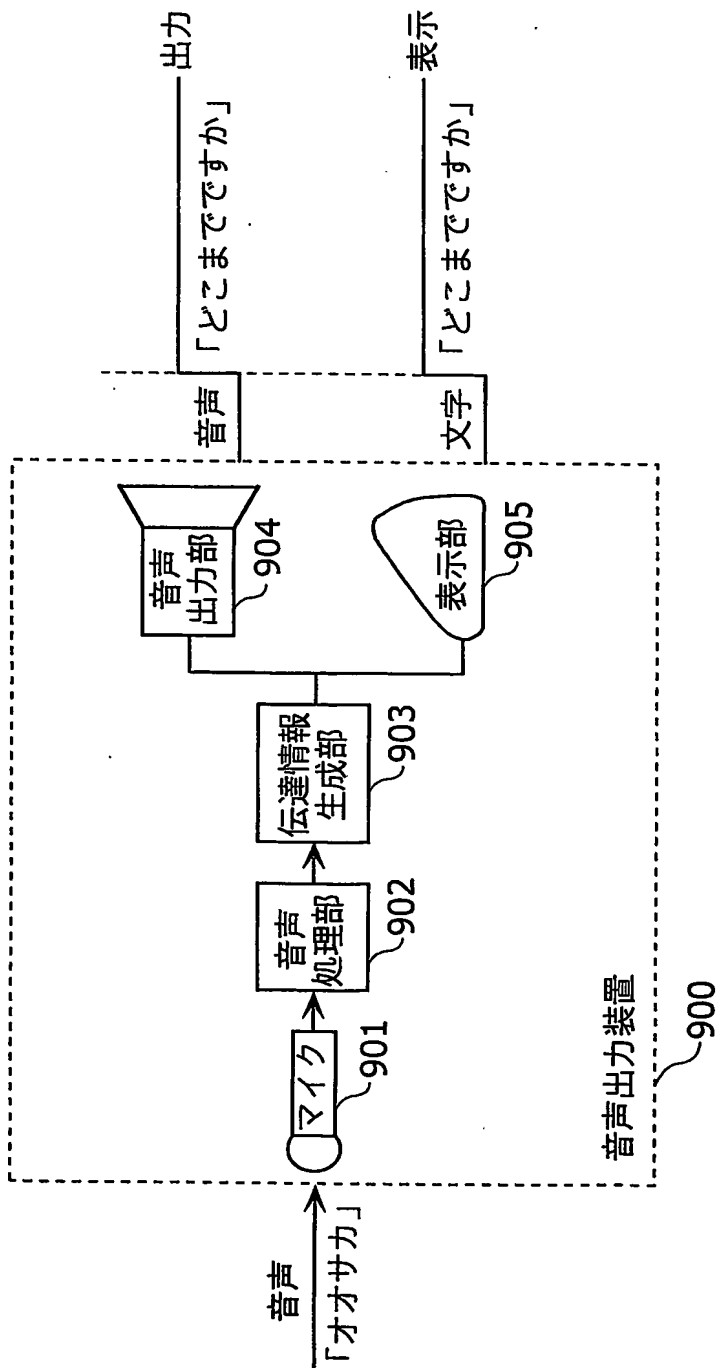


図2

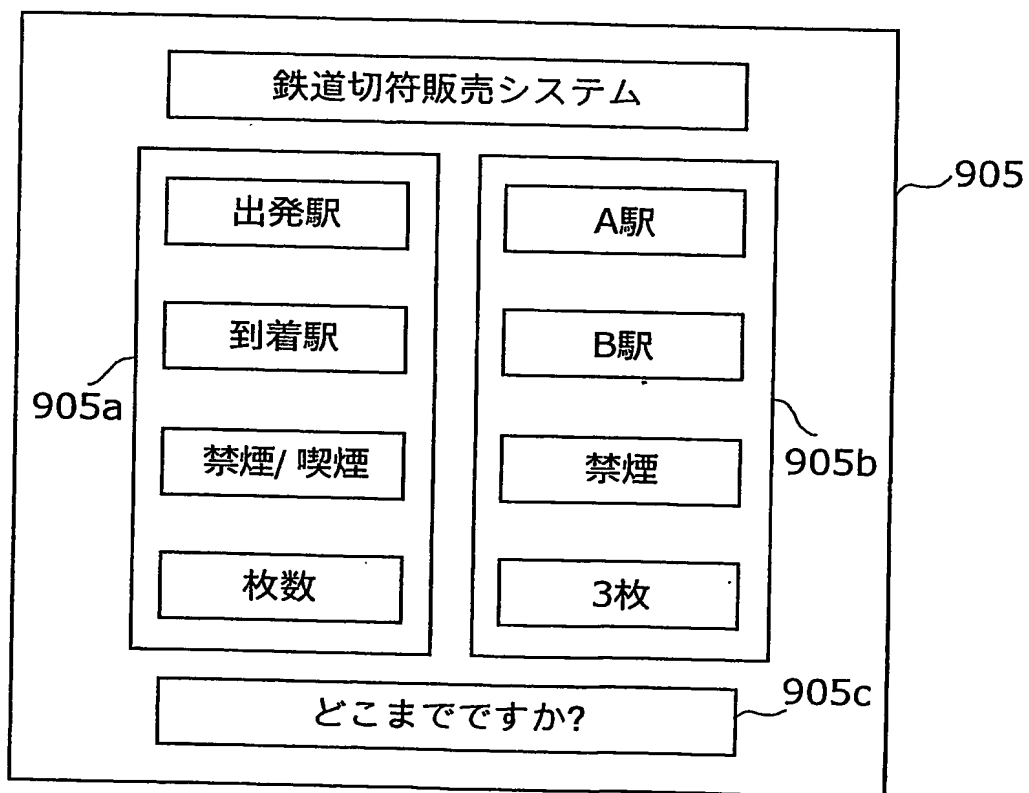


図3

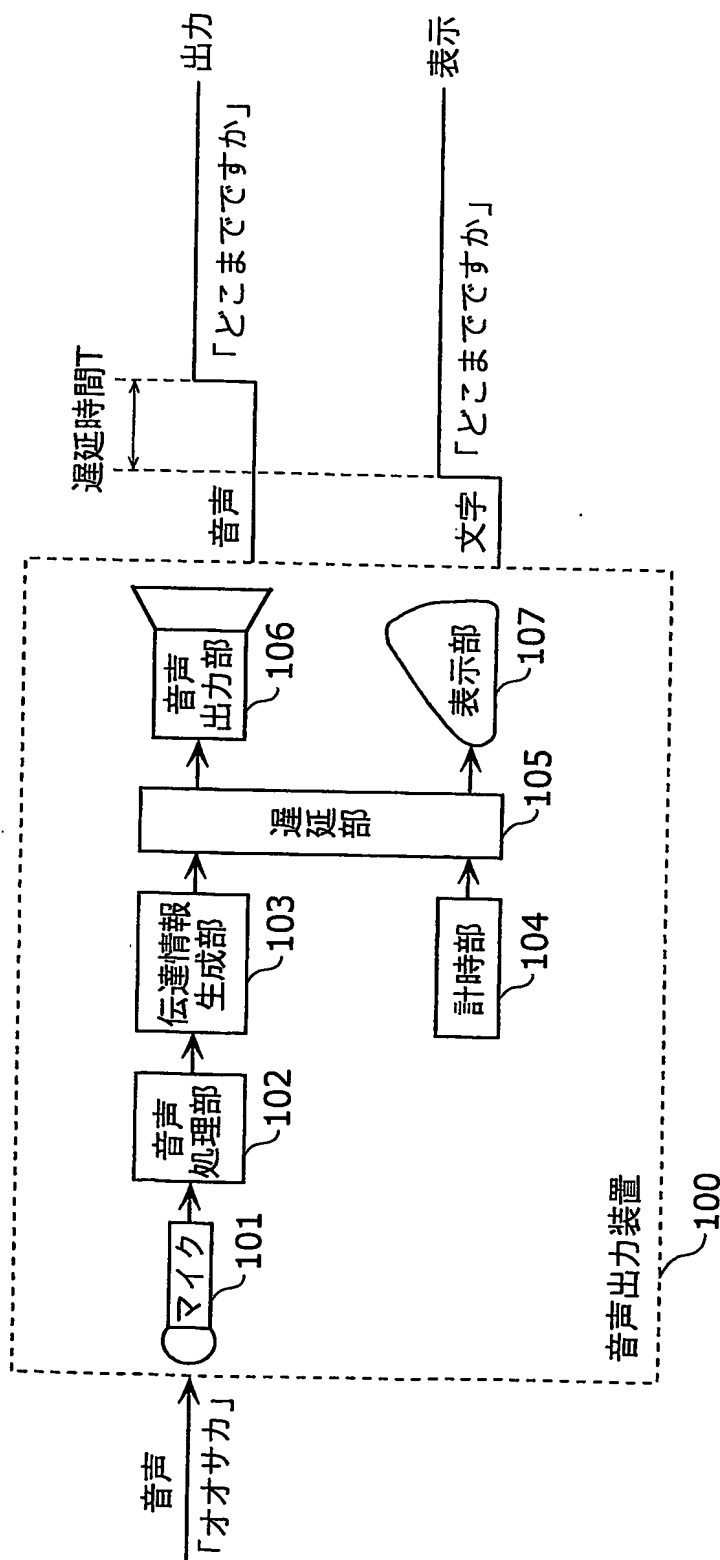


図4

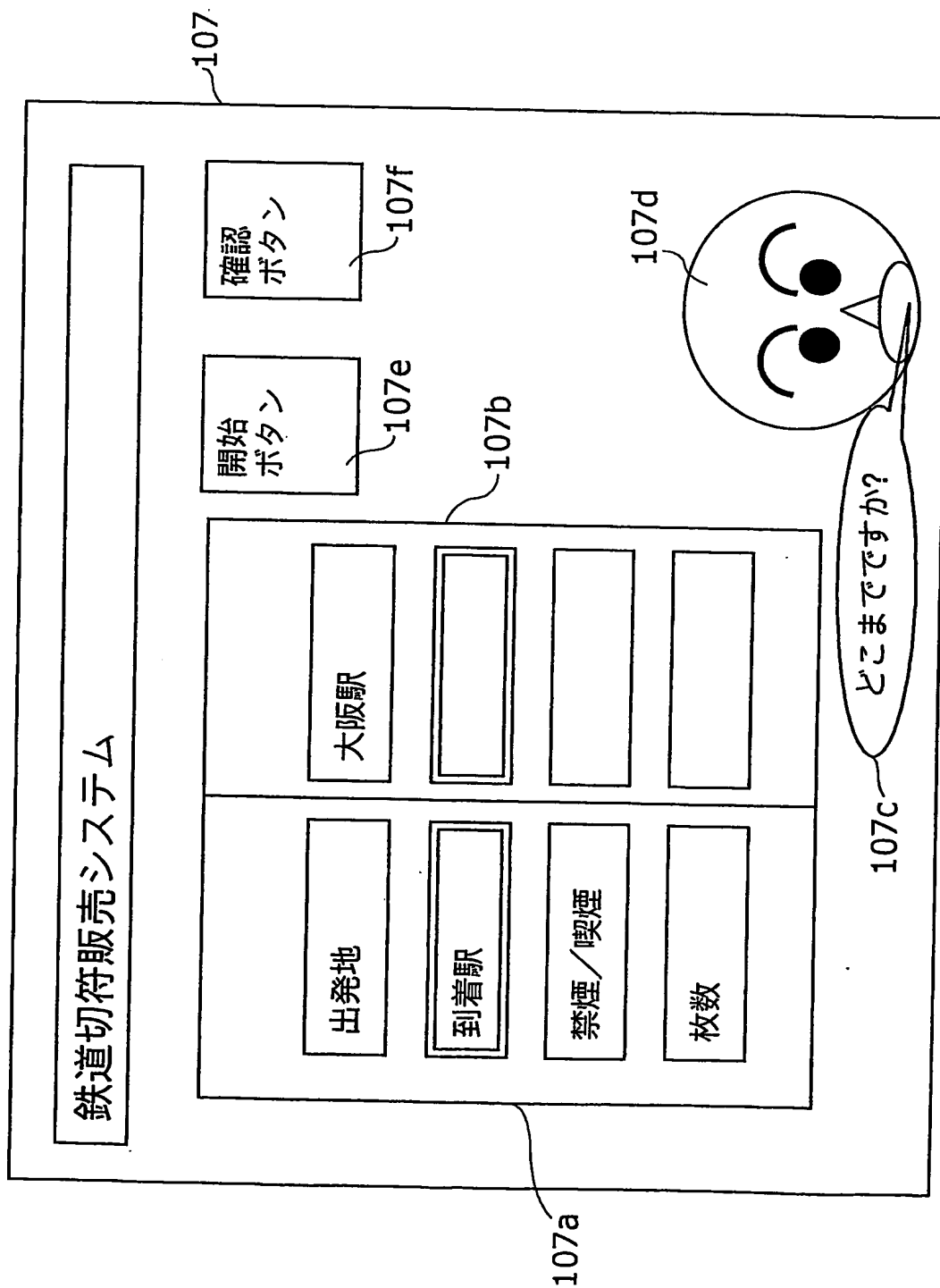


図5

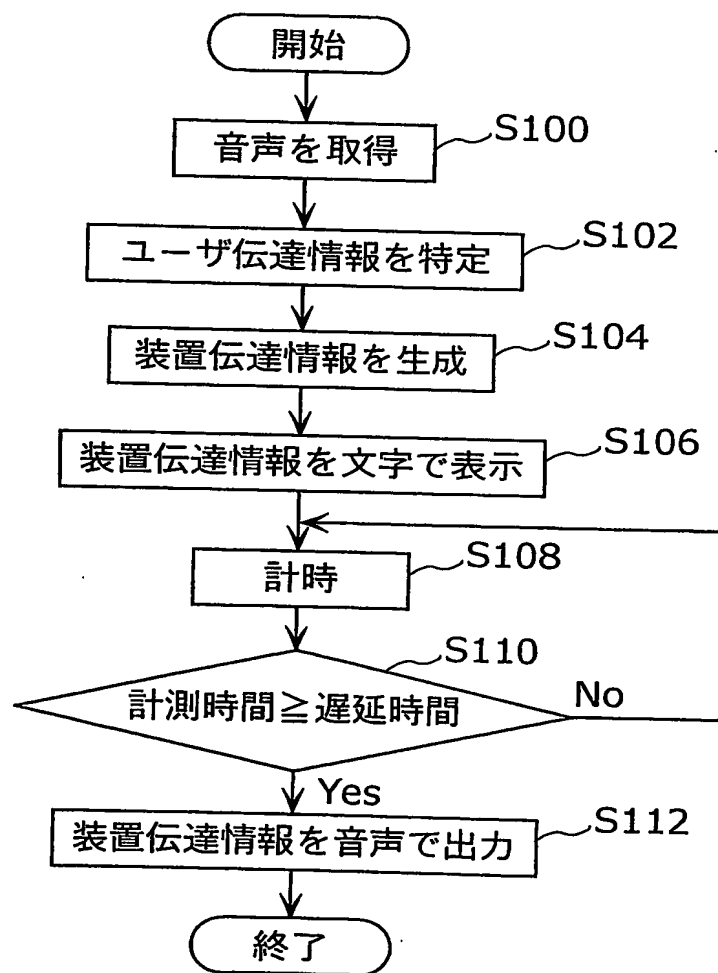


図6

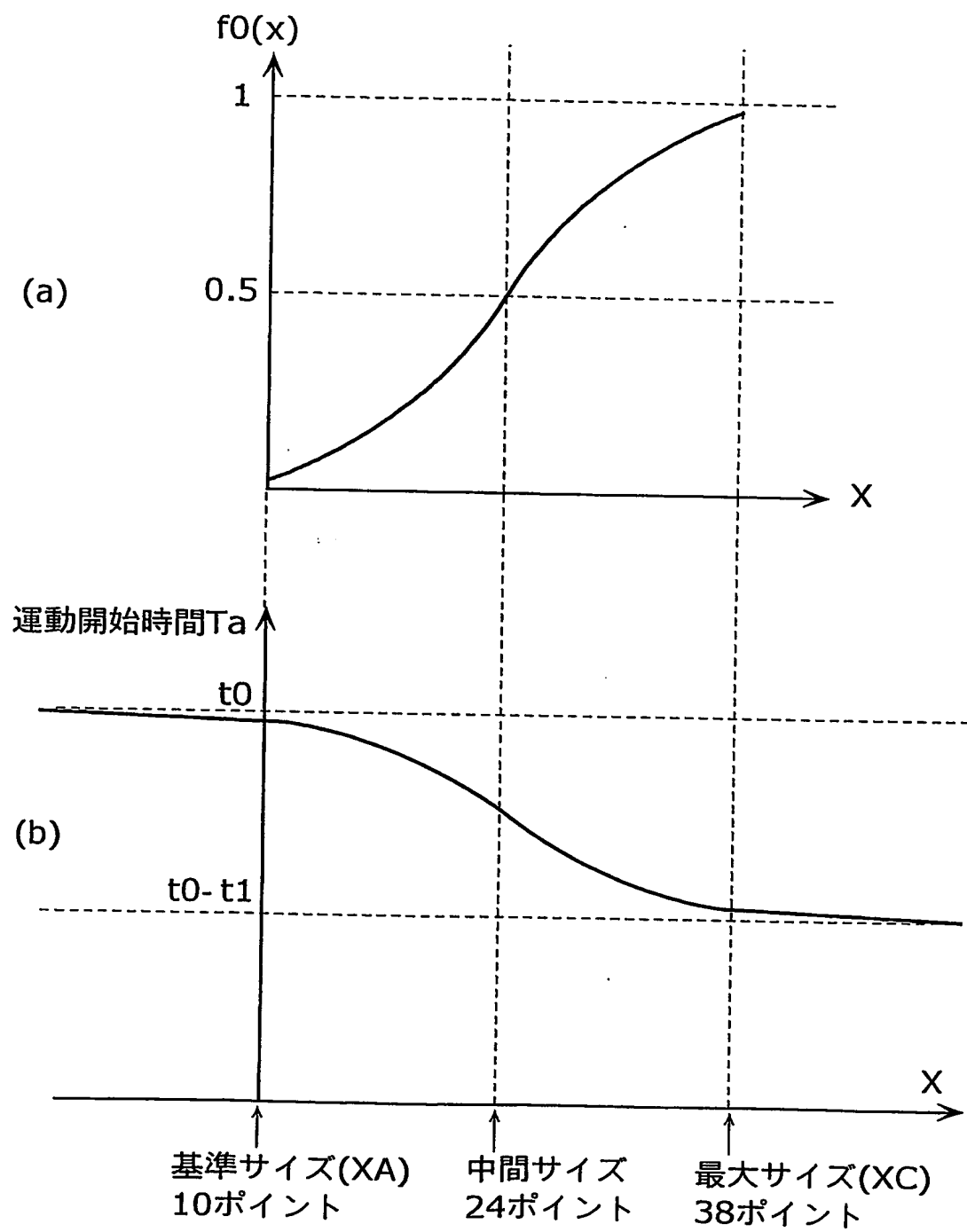


図7

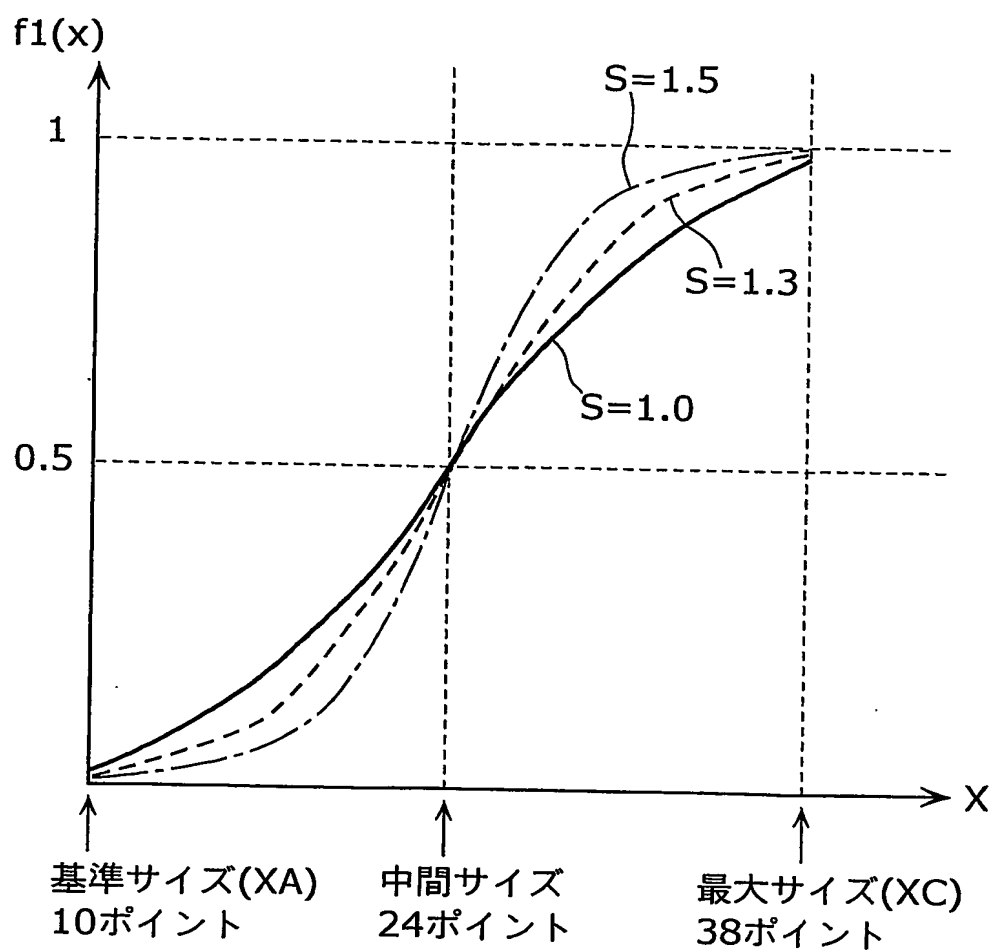


図8

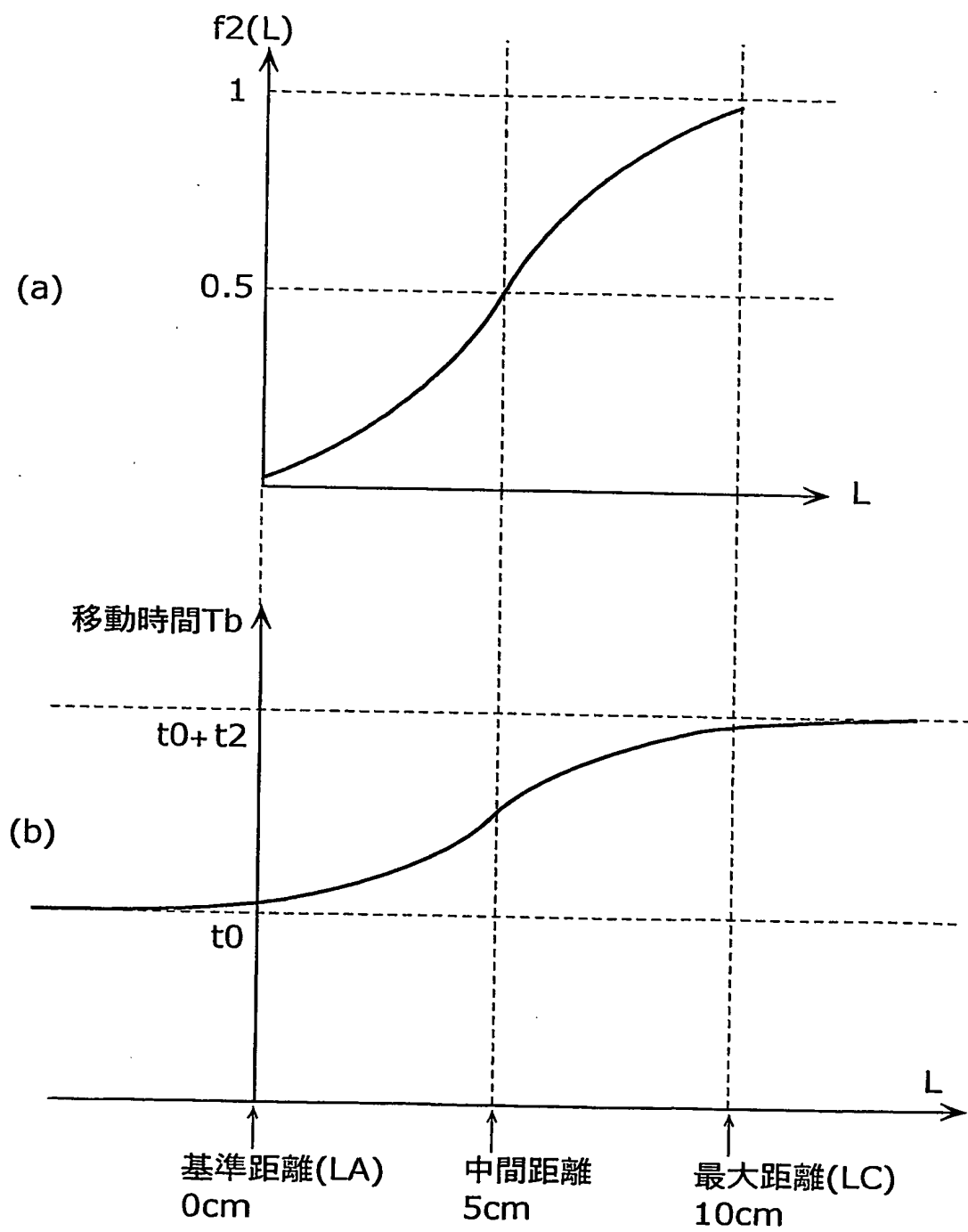


図9

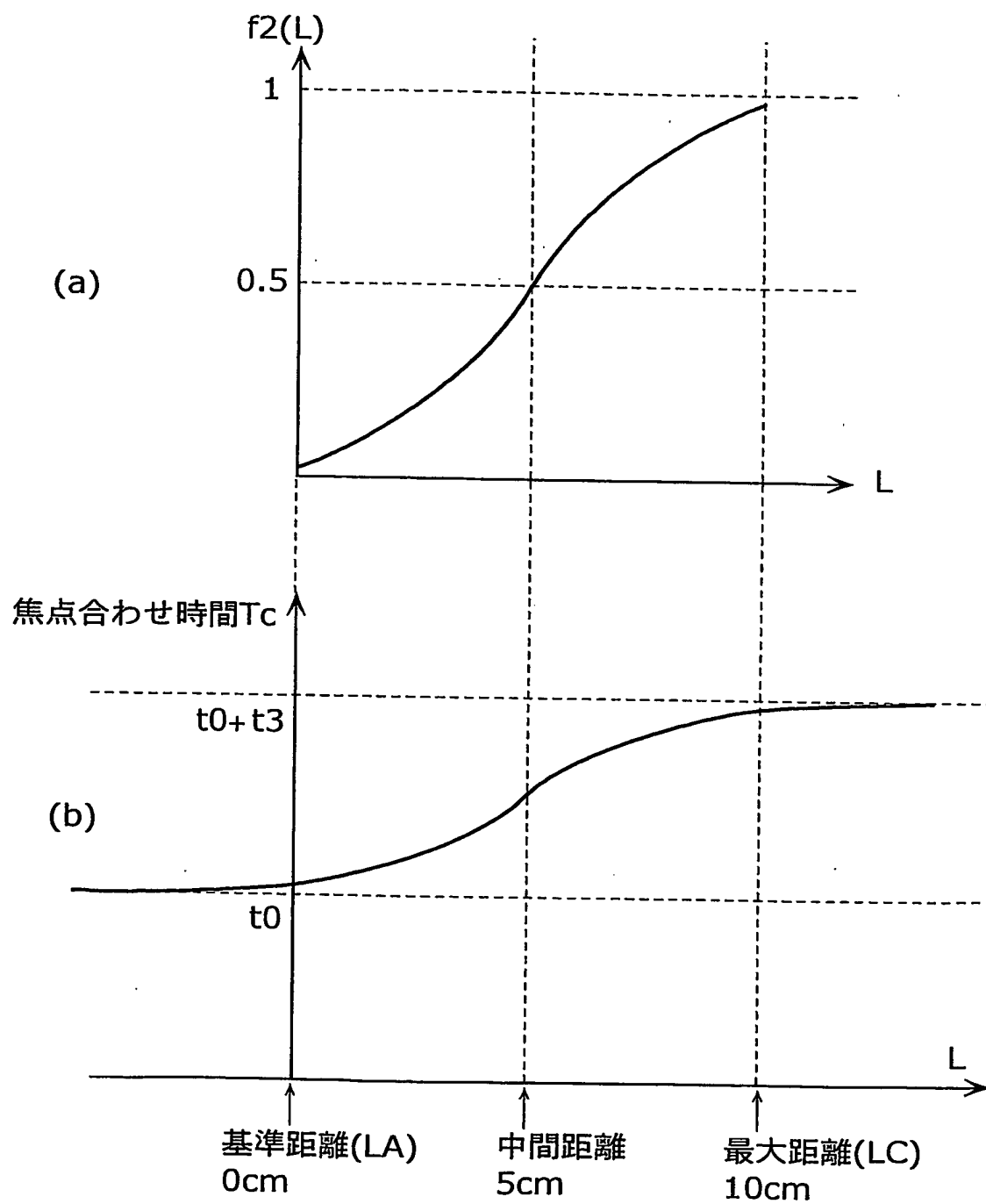


図10

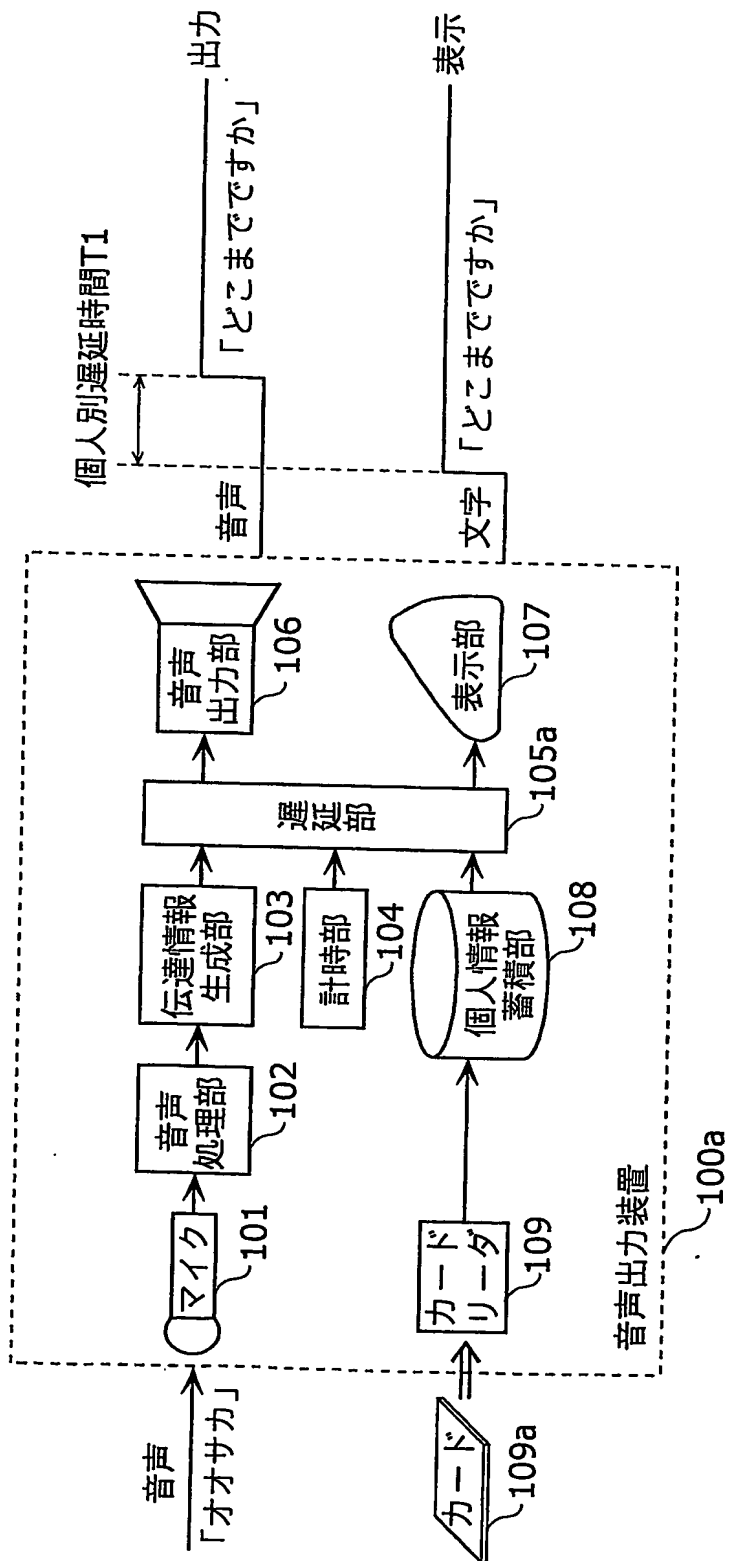


図11

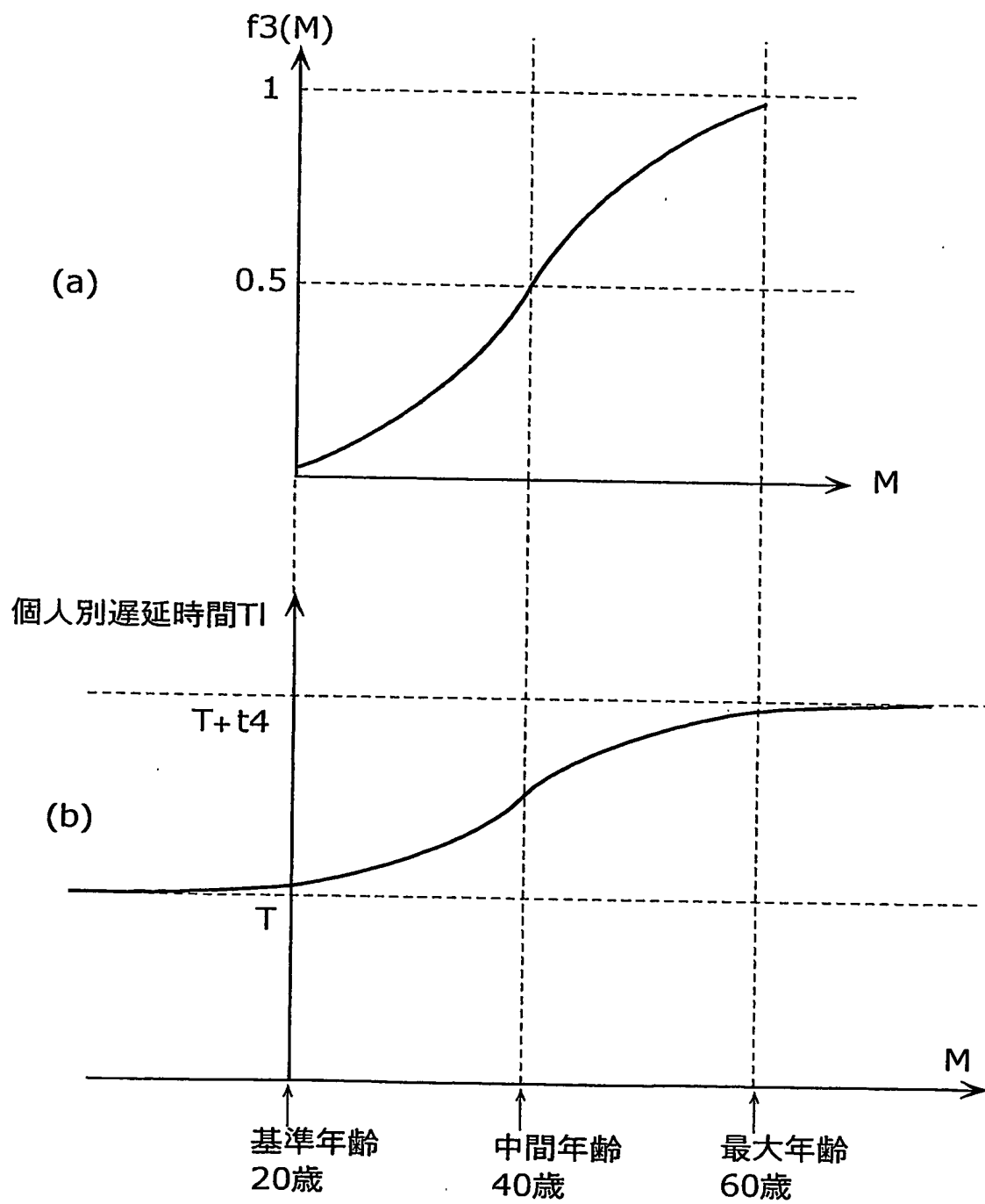


図12

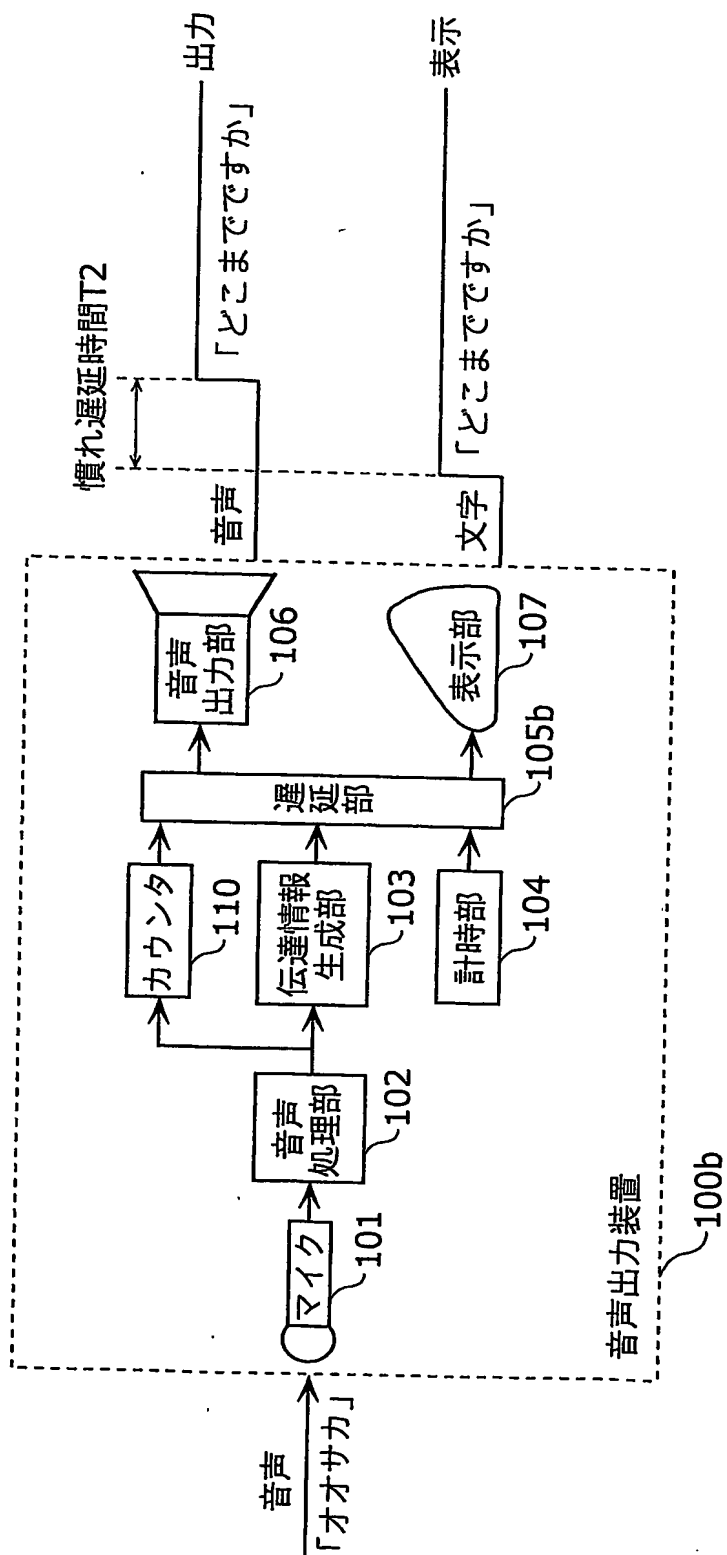
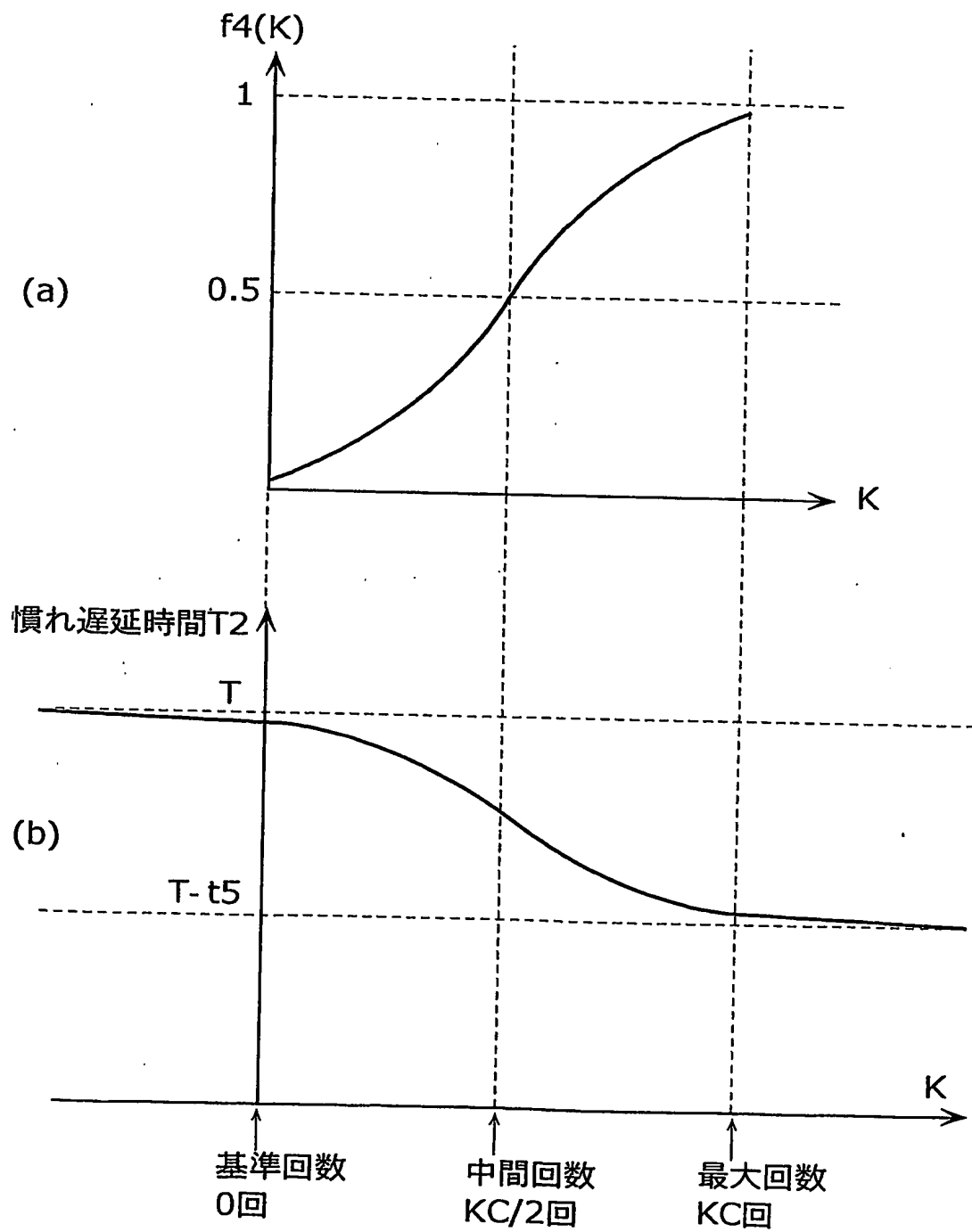


図13



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/006065

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ G10L13/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ G10L13/00, G10L21/06, G06F3/16

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2004
Kokai Jitsuyo Shinan Koho	1971-2004	Toroku Jitsuyo Shinan Koho	1994-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JTSPplus FILE (JOIS)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2003-108171 A (Clarion Co., Ltd.), 11 April, 2003 (11.04.03), Full text; all drawings (Family: none)	1, 23, 28
A	JP 6-321030 A (Yazaki Corp.), 22 November, 1994 (22.11.94), Full text; all drawings (Family: none)	1-28
A	JP 8-54894 A (Fuji Electric Co., Ltd.), 27 February, 1996 (27.02.96), Full text; all drawings (Family: none)	1-28

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

28 July, 2004 (28.07.04)

Date of mailing of the international search report

17 August, 2004 (17.08.04)

Name and mailing address of the ISA/

Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/006065

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2003-99092 A (Chuo Joho Kaihatsu Kabushiki Kaisha), 04 April, 2003 (04.04.03), Full text; all drawings (Family: none)	1-28
A	Hiroshi TAMURA, Human Interface, Ohmsha, Ltd., 30 May, 1998 (30.05.98), pages 32 to 33, 69 to 79, 212 to 213, 383 to 411	1-28

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ G10L13/00

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ G10L13/00, G10L21/06,
G06F3/16

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996
 日本国公開実用新案公報 1971-2004
 日本国実用新案登録公報 1996-2004
 日本国登録実用新案公報 1994-2004

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JTSPplusファイル (JOIS)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	JP 2003-108171 A (クラリオン株式会社) 2003. 04. 11, 全文, 全図 (ファミリーなし)	1, 23, 28
A	JP 6-321030 A (矢崎総業株式会社) 1994. 11. 22, 全文, 全図 (ファミリーなし)	1-28
A	JP 8-54894 A (富士通テン株式会社) 1996. 02. 27, 全文, 全図 (ファミリーなし)	1-28

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
 「O」 口頭による開示、使用、展示等に言及する文献
 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献
 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
 「&」 同一パテントファミリー文献

国際調査を完了した日

28. 07. 2004

国際調査報告の発送日

17. 8. 2004

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)
 郵便番号100-8915
 東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)
 山下 剛史

5C 3352

電話番号 03-3581-1101 内線 3541

C (続き) . 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	J P 2003-99092 A (中央情報開発株式会社) 2003. 04. 04, 全文, 全図 (ファミリーなし)	1-28
A	田村博, ヒューマンインターフェース, オーム社, 1998. 05. 30, p. 32-33, 69-79, 212-213, 383-411	1-28